

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 08:58:20 ; Search time 5973 Seconds
(without alignments)
11178.853 Million cell updates/sec

Title: US-10-070-386-1
Perfect score: 1378
Sequence: 1 gtcgacgtgggtgctgcatat.....accatcaatcaactaaca 1378

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1378	100.0	1378	6 BD013073	BD013073 Regulator
2	48.2	3.5	2000	6 AX655393	AX655393 Sequence
3	42.2	3.1	7218	6 T66494	T66494 Sequence 14
4	41.8	3.0	231585	2 AC103186	AC103186 Rattus no
5	41.8	3.0	329549	2 AC117331	AC117331 Rattus no
6	41.6	3.0	164377	2 AC136809	AC136809 Rattus no
7	41.6	3.0	226170	2 AC105687	AC105687 Rattus no
8	40	2.9	6197	6 AX344849	AX344849 Sequence
9	39.8	2.9	221649	2 AC137392	AC137392 Rattus no
10	39.8	2.9	262854	2 AC103016	AC103016 Rattus no
11	39.6	2.9	342084	2 AC128752	AC128752 Rattus no
12	39.6	2.9	81624	3 AC099763	AC099763 Caenorhab
13	39.6	2.9	290066	2 AC129421	AC129421 Rattus no
14	39.4	2.9	81921	10 AL805938	AL805938 Mouse DNA
15	39.4	2.9	186958	10 AC130210	AC130210 Mus muscu
16	39.4	2.9	195411	10 AL669958	AL669958 Mouse DNA
17	39.4	2.9	218807	2 BX511249	BX511249 Mus muscu
18	39	2.8	237119	10 AC115291	AC115291 Mus muscu
19	38.8	2.8	16512	3 CEU49795	CEU49795 Caenorhabd

C 20	38.8	2.8	34919	3 AF100663	AF100663 Caenorhab
C 21	38.8	2.8	139549	2 AC114644	AC114644 Mus muscu
C 22	38.6	2.8	129489	10 AL713860	AL713860 Mouse DNA
C 23	38.6	2.8	143148	10 AL607034	AL607034 Mouse DNA
C 24	38.6	2.8	201295	2 BX901879	BX901879 Danio rer
C 25	38.6	2.8	225552	2 BX897745	BX897745 Danio rer
C 26	38.6	2.8	233819	2 AC095098	AC095098 Rattus no
C 27	38.6	2.8	234873	2 AC137028	AC137028 Rattus no
C 28	38.6	2.8	240425	10 CNG57XOT	AL713839 Mus muscu
C 29	38.6	2.8	272563	2 AC095469	AC095469 Rattus no
C 30	38.4	2.8	195611	10 AC121971	AC121971 Mus muscu
C 31	38.4	2.8	334371	2 AC109571	AC109571 Rattus no
C 32	38.2	2.8	106259	2 AC151909	Continuation (4 of
C 33	38.2	2.8	144298	10 AC127682	AC127682 Mus muscu
C 34	38.2	2.8	180907	9 AC131157	AC131157 Homo sapi
C 35	38.2	2.8	202270	10 AC102241	AC102241 Mus muscu
C 36	38.2	2.8	253768	2 AC125994	AC125994 Rattus no
C 37	38.2	2.8	302921	2 AC090998	AC090998 Homo sapi
C 38	38	2.8	194370	2 AL645472	AL645472 Mus muscu
C 39	38	2.8	212543	5 BX663615	BX663615 Zabrafish
C 40	38	2.8	220638	10 AL645723	AL645723 Mouse DNA
C 41	38	2.8	257760	2 AC126638	AC126638 Rattus no
C 42	37.8	2.7	114999	9 AP003421	AP003421 Homo sapi
C 43	37.8	2.7	142148	9 AC025465	AC025465 Homo sapi
C 44	37.8	2.7	167912	9 AC060232	AC060232 Homo sapi
C 45	37.8	2.7	170066	9 AL161730	AL161730 Human DNA

ALIGNMENTS

RESULT 1
BD013073
LOCUS
DEFINITION
Regulatory sequences and expression system functional in mold fungi.

ACCESSION
BD013073
VERSION
BD013073.1 GI:22093262
KEYWORDS
WO 0118219-A/1.
SOURCE
unidentified
ORGANISM
unclassified.

REFERENCE
1 (bases 1 to 1378)
Watanabe, M. and Murakami, T.
Regulatory sequences and expression system functional in mold fungi
Patent: WO 0118219-A 1 15-MAR-2001;
JOURNAL
MEIJI SEIKA KAISHA LTD./MANABU WATANABE, TAKESHI MURAKAMI
COMMENT
OS Mycelia sterilia
PN WO 0118219-A/1
PD 15-MAR-2001
PF 07-SEP-2000 WO 2000P006104
PR 07-SEP-1999 JP 99P 252851
PI MANABU WATANABE, TAKESHI MURAKAMI
PC C12N15/80, C12N1/15, C12P21/00//((C12N15/80, C12R1:645), (C12N1/15, C12R1:645),
PC C12R1:645),
PC (C12P21/00, C12R1:645)
CC

FEATURES

source
FH Key Location/Qualifiers
1..1378
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 1378; DB 6; Length 1378;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCGACGTGGGTGATATCATGCTGTGTCCAAACTGTGTAGTACTACGAAT 60
1 GTCGACGTGGGTGATATCATGCTGTGTCCAAACTGTGTAGTACTACGAAT 60
DB 1 GTCGACGTGGGTGATATCATGCTGTGTCCAAACTGTGTAGTACTACGAAT 60

OY	6	GAGGAAAGAAACGGGCGGTGTGTGGCAGCTGAACACTGGAAGAGAGACCAAGAATTAATTC	120
Db	61	GAGGAAAGAAACGGGCGGTGTGTGTGGCAGCTGAAGACTGAAGAGAGAGACCAAGAATTAATTC	120
OY	121	ACAAATGCGATACGGTTGCATCAATGCTTGTTCAAAGAGACGTTGCATCTCCTGTGT	180
Db	121	ACAAATGCGATACGGTTGCATCAATGCTTGTTCAAAGAGAGACGTTGCATCTCCTGTGT	180
OY	181	TCCCTCTCTTGTTGTAACAAGTCAAGTATCGGATGACACCCACCCGACAGAAATCTTG	240
Db	181	TCCCTCTCTTGTTGTAACAAGTCAAGTATCGGATGACACCCACCCGACAGAAATCTTG	240
OY	241	GAGTTCAAAAGGGGGTGTCTACGGGCACTTAAGGATATAGATGGGATGAGGTTTACGTAA	300
Db	241	GAGTTCAAAAGGGGGTGTCTACGGGCACTTAAGGATATAGATGGGATGAGGTTTACGTAA	300
OY	301	GCTGAAAGCTATTTACGAGACATGAGACAAACGAAATAACAACGGTTGATGTCGTTCCGT	360
Db	301	GCTGAAAGCTATTTACGAGACATGAGACAAACGAAATAACAACGGTTGATGTCGTTCCGT	360
OY	361	GCTTACTAAAGTATATCAAGAGACAAACACCGGAAAGAAACCGATGCTGTGAGAGGG	420
Db	361	GCTTACTAAAGTATATCAAGAGACAAACACCGGAAAGAAACCGATGCTGTGAGAGGG	420
OY	421	TTCCTTTAGAGCTACATGATACGGGTGCAAGATATGAAACATCAAAATGGCCAAATCAAGTT	480
Db	421	TTCCTTTAGAGCTACATGATACGGGTGCAAGATATGAAACATCAAAATGGCCAAATCAAGTT	480
OY	481	AGTATACCTGACGCTACATCCGCTTCTCCGGATCTTGCCCTAAATAATATGTGCTGTCC	540
Db	481	AGTATACCTGACGCTACATCCGCTTCTCCGGATCTTGCCCTAAATAATATGTGCTGTCC	540
OY	541	GAACTGTCGGTACTGCTTCGTACTAACCTGTTCTTCCTGTTGAAGTCTTAGACAAACGCGCG	600
Db	541	GAACTGTCGGTACTGCTTCGTACTAACCTGTTCTTCCTGTTGAAGTCTTAGACAAACGCGCG	600
OY	601	CGTTTGTAAGCCTACATATGATGCAATCTTAAAGAGGAGACTGACACATTTCTTAAGGC	660
Db	601	CGTTTGTAAGCCTACATATGATGCAATCTTAAAGAGGAGACTGACACATTTCTTAAGGC	660
OY	661	ATCCATATAGGCAATTTGGGCGCTAAGTCCGGCACTTGGAAGAGATTAAGGGGGGTGTGAAGTG	720
Db	661	ATCCATATAGGCAATTTGGGCGCTAAGTCCGGCACTTGGAAGAGATTAAGGGGGGTGTGAAGTG	720
OY	721	GTTGTCAAAAGAGAGTGCATTTGGCTATACAGCCGCTAAGCAGGTGGGCTAGACGCTGT	780
Db	721	GTTGTCAAAAGAGAGTGCATTTGGCTATACAGCCGCTAAGCAGGTGGGCTAGACGCTGT	780
OY	781	CTGCACGCTGTGAATTAACGTCACTTGCTTAGGTATCCACTAATGTATGACAGATGCAAA	840
Db	781	CTGCACGCTGTGTAAACGTCACTTGCTTAGGTATGTCACTTAATGTATGACAGATGCAAA	840
OY	841	TGCTATTTGGGTTAAATGGGCACTAGTATGAGGTGCCGAAACACGTTTAGATCTAGTT	900
Db	841	TGCTATTTGGGTTAAATGGGCACTAGTATGAGGTGCCGAAACACGTTTAGATCTAGTT	900
OY	901	AAAGGGAAGCTGAAGCTGAACCTGTCAAAATAAGCCTGTTGGAATCAACGTTGATTA	960
Db	901	AAAGGGAAGCTGAAGCTGAACCTGTCAAAATAAGCCTGTTGGAATCAACGTTGATTA	960
OY	961	CCCAATTCAGTGTCAAGGGGTCTCTGATATGCTGTGAGCTTCCGTGTGCAATGTGGGGGT	1020
Db	961	CCCAATTCAGTGTCAAGGGGTCTCTGATATGCTGTGAGCTTCCGTGTGCAATGTGGGGGT	1020
OY	1021	AACATATTCATATGAGGAGAGAAATCAACTCAATTTCAATTTGAATCTAAACTATTTCTGG	1080
Db	1021	AACATATTCATATGAGGAGAGAGAAATCAACTCAATTTCAATTTGAATCTAAACTATTTCTGG	1080
OY	1081	GTAGAGATTTCTCAATGATCTTCTGCTGTCACTTACACATCAATCAATGAGGGGTCAACAAAGT	1140
Db	1081	GTAGAGATTTCTCAATGATCTTCTGCTGTCACTTACACATCAATCAATGAGGGGTCAACAAAGT	1140
OY	1141	ATACAGCTTCAATAGAGAGTGGCGCACTTGAATGACTACCGCATCGAACCCGGAACGGGTT	1200

Dd	1141	ATACAGTTCTATGAGAGTGGGCGATTGAAGTAGTACCCGATCGAACCCGGAAGCGTT	1200
Qy	1201	CAAGACATGGGCGTAGACTATCATAGAGTCATAGAAATAGAAAGAGCTTGAAGAAC	1260
Dd	1201	CAAGACATGGGCGTAGACTATCATAGAGTCATAGAAATAGAAAGAGCTTGAAGAAC	1260
Qy	1261	CATTCAATCTTAAGGCTCTCTTCTTTCTGCATCATCAAGAAATCATACCTCAAC	1320
Dd	1261	CATTCAATCTTAAGGCTCTCTTCTTTCTGCATCATCAAGAAATCATACCTCAAC	1320
Qy	1321	CAGGACCTCTTCTATCTTCCCTATAGCAATCCCAAAACCATCAATCAACCTTACA	1378
Dd	1321	CAGGACCTCTTCTATCTTCCCTATAGCAATCCCAAAACCATCAATCAACCTTACA	1378

RESULT 2			
AX655393			
LOCUS	AX655393	2000 bp	DNA
DEFINITION	Sequence 5263 from Patent WO03000898.		linear
ACCESSION	AX655393		
VERSION	AX655393.1	GI:29158207	

SOURCE
ORGANISM
Oryza sativa
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehretioideae; Oryzaceae; Oryza.

REFERENCE	AUTHORS	TITLE	JOURNAL
1	Chang, H.S., Chen, W., Cooper, B., Glarebrook, J., Goff, S.A., Hou, Y.M., Katagiri, F., Qian, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.	Plant genes involved in defense against pathogens	Parent: WO 03000898-A 5263 03-JUN-2003; Syngenta Participations AG (CH)

FEATURES	Location/Qualifiers
source	1..2000
	/organism="Oryza sativa"
	/mol_type="unassigned DNA"
	/db_xref="taxon:4530"
ORIGIN	

Query Match	3.5%	Score 48.2;	DB 6;	Length 2000;
-------------	------	-------------	-------	--------------

Best Local Similarity 10.0%; Pred.NO. 0.00/4;
Matches 79; Conservative 339; Mismatches 372; Indels 1; Gaps 1;

585 CCTAGGACAAGCGCCGCTTGTAGACCTACATGATGCCACATCTTAAGCAGGATCTG 644

b 28 SCARMSSRRMSKMGMSKYRKCS CGCKITTRKSKWYSASASGRITSKMSSGS YSG 87

645 AGACATTTCTTAAGGCATCCATATAGGCA TTGGCGCTAAGTCGGCATTTGAAGAGATTA 704

88 KGMKKRYKRSKWRGRGRGMRSSRMWGGYRRCARSGRMAGSGRMWGGSRMSYWM 147

705 GGGGGTGTGAAGTGTGTGTCAAAAGAGGTGATTTGGCTATACACCCGGCTTAACAG 764

148 CYARGCGSKRKKSKGSGWKTCRGARGSGSSGALKYKSGMSKRMNMSCGRSGCGR 207

765 GTGGCTAGCAGCTGTCTGCAGCTGTGATTAACGTACTTGGCTTAGGATATGCCACTAA 824

208 RSAISRYGTDRAKGYTKMTTYSASRCMAVMTTYSWACSSITWCRSRKRSMMVKKMR 267

825 TGTGACGAGATGCAAAATGCTGATTGGGTAAATAGGCATGTAGTGTAGCTGCCGGAAC 884

268 KMRWSRSYIGWISWSYKAWMCCLAYAKKSYISKWCIMIRGGGKCAURIWGRGISRWALAMIN 327

[illegible][illegible]

345 AATACACAGGCTGATATATCCCATTCAGTCATAGGAGGATCCCTCATATATCCATCCAC 20

388 ATACACATCCTATATATATATATATATATATATATATATATATATATATATATATATATAT 447

... ..

```

QY 1005 TGTGGATTGTGGGTAACCTATTTCATAGTGGGAGAAATGCAACTCTTAATTTCAATTGA 1064
DB 448 KEMASXYKMMRMRYRKRKKCSRTTWMGRTGSMGTGRCRYKKSGMKRCRWRMGR 507
QY 1065 ATCTAACTATTTCTGGTAGAGATTCTCATGTCTTTCGCTGCTGATCTTACACATCA 1124
DB 508 MYRMRRKRYMSARLYTRYCARRRKYSYSAARRKRCYRGRGYYMAGMMKRYRMYYMK 567
QY 1125 TGGGGGTCAACAAGTACGAGCTTCA-TAGAGAGTGGCGATTTGAAGTAGTACCGCAT 1183
DB 568 MMYRKRYSKSCWYCMSTYASGMSARAGAMCRSKMSAMSMSKMSRKRKCAASKR 627
QY 1184 CGAACCCGAGAGCGGTTCAAGACATGGCGGCTAGTAGATACATAGTAAACATTA 1243
DB 628 SSKAKYAMMGMTSGSRMSRWSYTCYMRKMSGMSKSTCMWYMSKYTAKYGSYWRYY 687
QY 1244 AAGAGAGCTTGAAGAACATTTCAATTCCTTGAAGGCTCTCTTCTTCTGATCAATCA 1303
DB 688 RAMCMYMMMYRRYRSYVTMYAMWYTSSTRMATGKMGYSGRWTSWYKCKCSWKYRSNW 747
QY 1304 GAATCATACACTCAACAGCAAGACTCTTCTTCTTCCCTATAGCAATTCGCAAAACCA 1363
DB 748 YTWMMWATTKMMRRYATRRMMMYRYSKMYTTCMTMGYMMWRTYMKRYMYKCT 807
QY 1364 TCAATCAACCT 1374
DB 808 KTYWYMSATYW 818

RESULT 3
LOCUS 166494 7218 bp DNA linear PAT 28-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION 166494
VERSION 166494.1 GI:2724471
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 7218)
Dorner, F., Schefflinger, F. and Falkner, F. Gunter.
TITLE
Recombinant fowlpox virus
JOURNAL
Patent: US 5670367-A 14 23-SEP-1997;
FEATURES
Location/Qualifiers
source
1..7218
/organism="unknown"
/mol_type="unasigned DNA"

ORIGIN
Query Match 3.1%; Score 42.2; DB 6; Length 7218;
Best Local Similarity 8.9%; Pred. NO. 0.53; 210; Indels 0; Gaps 0;
Matches 41; Conservative 212; Mismatches 210; Indels 0; Gaps 0;

QY 607 TAGACCTACATGATGCACATCTTAAAGCAGGATCTGAGACATTTCTAAGCATCCAT 666
DB 1450 TAGAAGAAATTTGTGTCRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1391
QY 667 ATAGCATTTGGCGGCTAAAGTCCGCATTTGAAGAGATAGAGGGGGTGTAAAGTGTGT 726
DB 1390 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1331
QY 727 CAAGAGGAGGTGATGCTATACCGCGCTAGACAGATGGGCTAGAGCGTGTGCGAC 786
DB 1330 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1271
QY 787 CTGTGAATAACGCTACTTGCTTAGATGATGCACTTAATGATGAGCATGAAATGCTGA 846
DB 1270 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1211
QY 847 TTGGGTTAAATGGGATATAGTGTAGTGGCGAAACACGTTTATAGATCTAGTTAAAGG 906
DB 1210 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1151

```

```

QY 907 AAGCTGAAGCTGAACTCTGAGAAATGAACCTGTTGGAATACAGCTTGATTAACCAAT 966
DB 1150 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1091
QY 967 TCAGTGTCAAGGATCTCGATATCTGAGCTTCCCTGCTGATTTGGCATTTGGGTAAT 1026
DB 1090 RRRRRRRRRRRRRRRRRRRRRRRRRRRATGCAAGTCCCTGACCTGAGCCAACTCGGA 1031
QY 1027 TTCATATGGGGAGAAATGCAACTTATTTTCAATGAATCTTA 1069
DB 1030 TTAATTTCTGTAGCGTATGCAACGAGAAATAATGTTATA 988

RESULT 4
LOCUS AC103186 231585 bp DNA linear HTG 13-MAY-2003
DEFINITION Rattus norvegicus clone CH230-107D17, *** SEQUENCING IN PROGRESS
ACCESSION AC103186
VERSION AC103186.5 GI:30578507
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 231585)
Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benhmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Casar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Crease, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Dlyva, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregiorgis, E., Geer, K., Gill, R., Girdly, M., Guerra, M., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,
Hollins, B., Howells, S., Hu, J., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, R., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshew, L., Louised, H., Lozano, R. J., Lu, X., Ma, U.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martine, E.,
Mawley, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nait, L.,
Nankervyl, C., Neal, D., Newton, N., Nguyen, A., Pal, S., Parks, K.,
Nwankweli, O., Okunolu, G., Olarnpusoon, A., Pal, S.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C.,
Plapper, F., Polindexter, A., Popovic, D., Prims, E., Pu, L.,
Puazo, M., Quintero, J., Rachlin, E., Reeves, K., Regier, M. A., Reish, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, R.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. U.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smjs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Soza, J.,
Steinle, M., Strong, R., Sutton, A., Svetek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejcek, Z., Uman, K.,
Vales, R., Vang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wiczek, R., Woden, H., Wozley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,

```

end_sequence: b7c35510...

B., Mapua, P., Martin, K.

Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, J., Morris, K., Morris, S., Mundaasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwokilemeh, O., Okwuon, G., Olarnpungoon, A., Pal, S., Parks, K.,
 Paternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plummer, F., Poidexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, J., Savatsky, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smajs, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabot, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K.,
 Valas, R., Wang, V., Villaseca, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 329549)
 Worley, K.C.
 Direct Submission
 Submitted (10-APR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 329549)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (09-OCT-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Sep 19, 2002 this sequence version replaced gi:21746236.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GUMC
 Center clone name: CH230-242A22
 ----- Summary Statistics
 Assembly program: Phrap, version 0.990329
 Consensus quality: 198199 bases at least Q40
 Consensus quality: 201002 bases at least Q30
 Consensus quality: 202294 bases at least Q20
 Estimated insert size: 210461; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence

* as soon as it is available and the accession number will
* be preserved.

1 38119: contig of 38119 bp in length

* 38120 38219: gap of unknown length

* 38220 291497: contig of 253278 bp in length

* 291498 291597: gap of unknown length

* 291598 294124: contig of 2577 bp in length

* 294125 294224: gap of unknown length

* 294225 329549: contig of 35325 bp in length.

FEATURES

source

1..329549

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-242A22"

/complement(2196..2863)

/note="Clone boundary

clone end:Spf

site:Mbol

end_sequence:RXAAPIITV"

38220..39946

/note="wgs_contig"

misc_feature

Query Match 3.0%; Score 41.8; DB 2; Length 329549;

Best Local Similarity 47.2%; Pred. No. 0.8; Matches 127; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

ORIGIN

1110 CACTTACGACATCATGTGGGGCTCAACAAGTTATCAGCTTCATTAGAGAGTGCGGCATTGA 1169

DB 278081 CACACACACACACACACACACACACACACACACAGAGAGAGAGAGAGAGAGAGAGA 278140

OY 1170 AGTAGCTACCCGATCGAACCCGGAGCGGTTCAAGACATGGCGTAAGTATACATAGA 1229

DB 278141 CAGAGAGACAGAGAGACAGAGAGACAGAGAGAGACAGACAGAGAGACAGAGAGAGA 278200

OY 1230 GTCATAGAAACATATAAGAGAGCTTTAAGAACCATTCCTTAAGGGCTCTCTCTTT 1289

DB 278201 GGCGAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 278260

OY 1290 CTGATCATCATCAAGATCATCTACCTCAACACGAGAGACTCTTATCTCTCCCTATAGA 1349

DB 278261 CTGAGAGACATTTTAAAAACATCTCACACAGTAAAATTAAACCAATCCAGAAATATATA 278320

OY 1350 ATTCCCAAAACCATCAATCAACCTTACA 1378

DB 278321 CTTTCAATAAATCTTCACTCACTCAATAAA 278349

RESULT 6

AC136809 164377 bp DNA linear HTG 23-NOV-2002

LOCUS AC136809

DEFINITION Rattus norvegicus clone CH230-349D12, *** SEQUENCING IN PROGRESS

ACCESSION AC136809

VERSION AC136809.2 GI:25188374

KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 164377)

Murphy,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Angilano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Bernhmed,F.,
Bialwalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Byrant,N., Buhan,C., Burck,P., Butrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Claveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dedertch,D.,

[illegible]

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

3 (bases 1 to 226170)
Rat Genome Sequencing Consortium.
Submitted (08-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Sep 18, 2002 this sequence version replaced gi:21743897.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rac/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GNNC
Center clone name: CH230-1111
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 19816 bases at least Q40
Consensus quality: 202314 bases at least Q30
Consensus quality: 204704 bases at least Q20
Estimated insert size: 229286; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_delta.html).
* NOTE: This is a "working draft" sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 7584: contig of 7584 bp in length
*
* 7585 7684: gap of unknown length
*
* 7685 224621: contig of 216537 bp in length
*
* 224622 224721: gap of unknown length
*
* 224722 226170: contig of 1449 bp in length.
*
Location/Qualifiers
1. 226170
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-1111"
45780..46902
/note="wgs_contig"

ORIGIN

Query Match 3.0%; Score 41.6; DB 2; Length 226170;
Best Local Similarity 49.1%; Prid. No. 0.91;
Matches 110; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

1155 AGAGTGGCGATTGAAGTACTGATCCGATCGAACCCGGAACGGCTTCAGACATCGGCGG 1214
|||||
Db 61562 AGATGCGCTGGAATTTAAGAGATACAAAGAGAAAGGGGAGACATTATGAAGTCAACG 61507
|||||
Qy 1215 ACGTAGATAGATAGAGTCATAGAAACATTAAGAGAGCTTGAAAGACATTCAATCTTA 1274
|||||
Db 61502 ACATGATATGACACATTTACTAGTACGTTTCACAAAGCATCATTTCTATTAAATATTAT 61444
|||||
Qy 1275 GGCTCTCTCTTTCTTTCGATCATCATCAAGATCATACACTCAACACGGAATCTTTTCT 1334

individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the end of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: G1YH

Center clone name: CH230-171L19

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 213826 bases at least Q40

Consensus quality: 217363 bases at least Q30

Consensus quality: 219200 bases at least Q20

Estimated insert size: 225374; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 251707: contig of 251707 bp in length
 * 251708 251807: gap of unknown length
 * 251808 253055: contig of 1248 bp in length
 * 253056 253155: gap of unknown length
 * 253156 254459: contig of 1304 bp in length
 * 254460 254559: gap of unknown length
 * 254560 256577: contig of 2018 bp in length
 * 256578 256677: gap of unknown length
 * 256678 262254: contig of 5577 bp in length.

FEATURES

source Location/Qualifiers

1. 262254

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-171L19"

478. 1118

misc_feature /note="clone boundary"

clone end:17

site:ECORI

end sequence: BH288239"

misc_feature complement (248949..249678)

/note="clone boundary"

clone end:5p6

site:ECORI

end sequence: BH288241"

ORIGIN

Query Match

Best Local Similarity 50.8%; Prid. No. 3.2;

Matches 95; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

11 GTGTGATATCATGCGTGTGCTGCTCAAACTGTGTAGTACTACGATGAGAGAAAGAA 70

DB 31823 GTGTGGATACCTCTCTTTGTTAGATATGTTAAAGAAACATGAAAAAGCTCAGAG 31882

QY 71 ACGGTGATGTTGGAGCTGAGCTGAGAGCTGAAGAGAGCCAAAGATATATTCACATGGCAT 130

DB 31883 AAGGCTTAGCAAGGCTTCTTACAAAGAGAAAGCTTTCAACAGAGAGAGAAAAAGCTGA 31942

QY 131 ACGGTGATCAATGCTTGTTCAGAGAGACGTTGATCTACCTGCTGTTCCCTCTTC 190

DB 31943 AAGCTGATCAAAACAGTCTCTCCATCCTCCCTGTGTTCCATCTGCTTCCCTCTCT 32002

QY 191 GTTGTAC 197

DB 32003 GTTGCAC 32009

RESULT 11

AC128752

LOCUS

DEFINITION

AC128752.3

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 342084)

REFERENCE

AUTHORS

Muzny,D.,Marie, Metzker,M.,Lee, A.,Abramson,S., Adams,C., Alder,J.,

Allen,C., Allen,H., Albrooks,S., Amin,A., Angiano,D.,

Anylebech,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,

Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benhammed,F.,

Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,

Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,

Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,

Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,

Davila,M.L., Davis,C., Day-Carroll,L., De Anda,C., Dederich,D.,

Delgado,O., Dawson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,

Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,

Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,

Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,

Gebregregorys,E., Geer,K., Gill,R., Grady,M., Guerra,M., Guevara,W.,

Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,

Harvey,Y., Havlik,P., Hawes,A., Henderson,N., Hernandez,J.,

Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,

Hollins,B., Howells,S., Huijck,S., Hume,J., Idlebird,D., Jackson,A.,

Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,

Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,

Kowals,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,

Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,

Lorenshuew,L., Louised,H., Lozano,R.J., Lu,X., Ma,D.,

Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A.,

Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,

Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,

Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,

Morgan,M., Morris,K., Morris,S., Mundasa,M., Murphy,M., Nair,L.,

Nankervis,C., Neal,D., Newton,N., Nguyen,S., Norris,S.,

Nwackemele,O., Okunodu,G., Olarnunsgoon,A., Pal,S., Parks,K.,

Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankuch,C.,

Plopper,F., Poindexter,A., Popovic,D., Primm,E., Pu,L., L.,

Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,

Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,

Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,

Sherry,M., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,

Shetty,J., Shwartsbeyn,A., Sison,I., Sitter,C.D., Smajls,D.,

Sneed,A., Sodergren,B., Song,X.-Z., Sorelle,R., Soza,D.,

Steinle,M., Strong,R., Sutton,A., Svatek,A., Taboi,P., Taylor,C.,

Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umanan,K.,

Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,

Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,

Williams,G., Willson,R., Wleczek,K., Woden,H., Worley,K.,

Wright,D., Wright,R., Wu,J., Yakub,S., Yen,U., Yoon,L., Yoon,V.,

Yu,F., Zhang,J., Zhou,J., Zhou,J., Zhou,S., Zhao,S., Dunn,D., von

Weinstock,G., A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,

Direct Submission

Unpublished

2 (bases 1 to 342084)

TITLE

JOURNAL

REFERENCE

REFERENCE	AUTHORS	JOURNAL	TITLE
3	Morley, K.C.	Direct Submission	Submitted (24-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3	(bases 1 to 342084)	Rat Genome Sequencing Consortium.	Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3	On Nov 15, 2002 this sequence version replaced gi:23803473.	The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contig will be indicated in the feature table.	
3	----- Genome Center	Center: Baylor College of Medicine	Center code: BCM
3	Web site: http://www.hgsc.bcm.tmc.edu/	Contact: hgsc-help@bcm.tmc.edu	----- Project Information
3	Center project name: KALG	Center clone name: CH230-466L2	----- Summary Statistics
3	Assembly program: Phrap; version 0.990329	Consensus quality: 14387 bases at least Q40	Consensus quality: 144547 bases at least Q30
3	Consensus quality: 144951 bases at least Q20	Estimated insert size: 177730; sum-of-contigs estimation	Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
3	-----	* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/doc/Genbank_draft_data.html).	* NOTE: This is a 'working draft' sequence. It currently
3	* consists of 1 contigs. Gaps between the contigs	* are represented as runs of N. The order of the pieces	* is believed to be correct as given, however the sizes
3	* of the gaps between them are based on estimates that have	* provided by the submitter.	* This sequence will be replaced
3	* by the finished sequence as soon as it is available and	* the accession number will be preserved.	1 342084: contig of 342084 bp in length.
3	Location/Qualifiers	1. 342084	/organism="Rattus norvegicus"
3	/mol_type="genomic DNA"	/db_xref="taxon:10116"	/clone="CH230-466L2"
3	184048..186150	/note="wgs_contig"	
3	ORIGIN	Query Match	2.9%; Score 39.8; DB 2; Length 342084;
3	Best Local Similarity	50.8%; Pred. No. 3.3;	Matches 95; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
3	11 GTGATGATATCATGAGCTGTGGTCTCAAACTGTGTTAGTACTACGAATGAGAAAGAA 70	DB	90802 GTGTTGGATCTCTCTTTGTAGAAATGTGTTAAAGAACATGAAAAAGTCTCAGAG 90861
3	71 ACGGTGCTGTGTGAGCAGCTGAAGACCTGAAGAAAGAGCCAAATATTAATTCACAAATCCGAT 130	DB	90862 AAGGCTTAGACAAAGGCTTTTACAAAGAAAGGCTTTCAACAAAGAAAGGAAAAAGCTGA 909211

QY	131	ACGGTTGCATCAATGCTGTGTTCAAGAGACGCTGATCTACCTGGTGTCCCTCTTC	190
Db	90922	AAGCTGAATCAAACTGTTCTCCACCTCCCTTGCTGTTCATCTCTCTCTCTCT	90988
QY	191	GTGTGTAC 197	
Db	90982	GTTCGAC 90988	
RESULT 12			
AC099763/c	AC099763	81624 bp DNA linear	INV 20-NOV-2001
DEFINITION	Caenorhabditis briggsae cosmid CB022010, complete sequence.		
LOCUS	AC099763		
VERSION	AC099763.1	GI:17017635	
KEYWORDS	HTG.		
SOURCE	Caenorhabditis briggsae		
ORGANISM	Caenorhabditis briggsae		
REFERENCE	Eukaryote; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.		
AUTHORS	1 (bases 1 to 81624)		
TITLE	Washington University Genome Sequencing Center.		
JOURNAL	The C. briggsae Genome Sequencing Project		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 81624)		
TITLE	Waterston, R.		
JOURNAL	Direct Submission		
REFERENCE	Unpublished		
AUTHORS	3 (bases 1 to 81624)		
TITLE	Waterston, R.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (20-NOV-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA		
JOURNAL	Submitted by:		
COMMENT	Genome Sequencing Center Department of Genetics, Washington University St. Louis, MO 63110, USA email: tw@nematode.wustl.edu		
NOTICE:	This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.		
FEATURES	<p>This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m13 subclone.</p> <p>Location/Qualifiers</p> <p>1..81624</p> <p>/organism="Caenorhabditis briggsae"</p> <p>/mol_type="genomic DNA"</p> <p>/strain="Gujarat G16"</p> <p>/db_xref="taxon:6238"</p>		
ORIGIN	<p>Query Match 2.9%; Score 39.6; DB 3; Length 81624;</p> <p>Best Local Similarity 58.5%; Pred. No. 3.6;</p> <p>Matches 69; Conservative 0; Mismatches 49; Indels 0; Gaps 0;</p>		
QY	34	TCCAAACTGTGTAGTACTGACTGATGAGAAAGAAACGCTGTGTGCGACGCTGAA	93
Db	8467	TTCAAAAATCTTATTATTATTCGACGTGTTAAAAAAGAAAGTTTGTACCAATTGAG	8408
QY	94	GACGGAAGAGAGCCAAAGATAATTCACATCGATACGTTGCATCAATGCTGTT	151
Db	8407	ACATGAATAATCAGTAAAAAATTAATATGAGTAAAGCAAAAGTACAACTGTTAGTT	8350

RESULT 13 AC129421/c	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
AC129421	290066 bp	DNA	linear	HTG 20-NOV-2002					
Rattus norvegicus clone CH230-62C13, ***	SEQUENCING IN PROGRESS								
AC129421	290066 bp	DNA	linear	HTG 20-NOV-2002					
AC129421.3	GI:25138103	HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.							
Rattus norvegicus (Norway rat)									
Rattus norvegicus									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.									
1 (bases 1 to 290066)									
Muzny, D. Marie, Metzger, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D., Anyalebech, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregregis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lounsbury, L., Loulsegod, H., Lozano, R. J., Lu, X., Ma, Z., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackeleneh, O., Okunolu, G., Olamunsgoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Plankkoc, C., Plapper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Puato, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Sma's, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabors, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umanai, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Woodan, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhao, S., Zhou, D., von Niederhausen, A., Weis, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G., and Gibbs, R. A.									
Direct Submission									
Unpublished									
2 (bases 1 to 290066)									
Worley, K. C.									
Direct Submission									
Submitted (30-JUL-2002)									
Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA									
3 (bases 1 to 290066)									
Rat Genome Sequencing Consortium.									

TITLE
Direct Submission

JOURNAL
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT
On Nov 20, 2002 this sequence version replaced gi:23267370. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GURV

Center clone name: CH230-62C13

Assembly program: Phrap; version 0.990329

Consensus quality: 216696 bases at least Q40

Consensus quality: 216655 bases at least Q30

Consensus quality: 221692 bases at least Q20

Estimated insert size: 216915; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently

consists of 5 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 40609: contig of 40609 bp in length

40610 40709: gap of unknown length

281950: contig of 241241 bp in length

281951 282050: gap of unknown length

282051 283370: contig of 1320 bp in length

283371 283470: gap of unknown length

283471 286076: contig of 2606 bp in length

286077 286176: gap of unknown length

286177 290066: contig of 3890 bp in length.

Location/Qualifiers

1.290066

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-62C13"

32547. 33462

/note="clone_boundary"

clone_end:5p6

site:

end sequence:BH287162"

40710. 42036

/note="wgs_contig"

280810. 281950

/note="wgs_contig"

ORIGIN

Query Match 2.9%; Score 39.6; DB 2; Length 290066;

Best Local Similarity 52.4%; Pred. No. 3 7;

Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

ORIGIN

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.choi.org>) from male C57Bl/60 mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.choi.org>

NEIGHBORING SOURCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES**Source**

1. 186858
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="7"
/map="7"

/clone="RP24-344N22"
/clone_lib="RPCI-24"

1334_1484

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B4"

repeat_region
/rpt_family="B4"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region
/rpt_family="B2"

repeat_region

14479..14860

repeat_region
/rpt_family="L1"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

repeat_region
/rpt_family="Alu"

Query Match 2.9%; Score 39.4; DB 10; Length 186858;
Best Local Similarity 54.5%; Pred. No. 4.2;

	Matches	79;	Conservative	0;	Mismatches	66;	Indels	0;	Gaps	0;
QY	1215	ACGTAGATACATAGAGTCTATAGAAACATTAAGAGAGCTTGAAAGACCATTCAAATCCTTA								1274
Db	130261	ATGTACAAACATATATCTCACACACAAATTAAGTAAATTGAAAAAGATTAATATCT								130202
QY	1275	GGGTCTCTCTTCTTCTGATCAGATCAAGAAATCATACACTCAAAACGAGAACTTTCT								1334
Db	130201	TAAATCAAAACACACAGGTACACACACAAACATTAACAAAAACAAAAACA								130142
QY	1335	ATCTTCCTTAAGCAATTCCAAAA								1359
Db	130141	ACCCTCCCTAAACAAAAA								130117

Search completed: February 9, 2005, 12:28:20
 Job time : 5980 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 08:54:14 ; Search time 752 Seconds
(without alignments)
10847.611 Million cell updates/sec

Title: US-10-070-386-1

Perfect score: 1378

Sequence: 1 gtgcagctgggtggtgatac.....accatcaatcaactaaca 1378

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1378	100.0	1378	5	AAf79834
2	48.2	3.5	2000	8	ADA71938
3	40	2.9	6197	6	ABN80257
4	37.2	2.7	2000	8	ADA71938
5	36.2	2.6	2672	6	ADA29004
6	36.2	2.6	2672	12	ADN96821
7	36.2	2.6	4590	5	AAH24065
8	35.8	2.6	134738	11	ACN44182
9	35.4	2.6	993	4	AAf81969
10	35.4	2.6	993	10	ADBE29362
11	35	2.5	6971	6	ABL33237
12	35	2.5	207433	5	ABZ72040
13	35	2.5	207433	8	ABX74891
14	35	2.5	207433	12	ADJ36614
15	35	2.5	207433	12	ADL81193
16	34.8	2.5	4416	13	ADS61188
17	34.6	2.5	5059	2	AAH84332
18	34.4	2.5	2799	10	ABV74497
19	34.4	2.5	8805	6	ABK40016
20	34.4	2.5	32874	9	ADA02648

21	34.4	2.5	32874	10	ADB72386	ADB72386 Human TBX
22	34.4	2.5	32874	10	ADE95896	ADE95896 Human TBX
23	34.2	2.5	1500	10	ADC92243	ADC92243 E. faeciu
24	34.2	2.5	90435	12	ADO59524	ADO59524 Human can
25	34	2.5	6283	6	ABK39991	ABK39991 Human che
26	34	2.5	6283	6	ABL32834	ABL32834 Human lmm
27	34	2.5	339234	12	ADQ59437	ADQ59437 Human can
28	33.8	2.5	648	10	ADK56087	ADK56087 Plant DNA
29	33.8	2.5	980	10	ADK59677	ADK59677 Plant DNA
30	33.8	2.5	103471	12	ADO97668	ADO97668 Mouse can
31	33.6	2.4	25179	13	ABD33248	ABD33248 Murine ca
32	33.6	2.4	191284	12	ADQ97957	ADQ97957 Mouse can
33	33.4	2.4	403	4	AAI36496	AAI36496 Probe #51
34	33.4	2.4	403	4	ABA26537	ABA26537 Probe #50
35	33.4	2.4	1781	2	AAV35130	AAV35130 Mouse WRN
36	33.4	2.4	4206	2	AAV35115	AAV35115 Mouse WRN
37	33.4	2.4	4792	2	AAH83004	AAH83004 Mouse WRN
38	33.4	2.4	5058	2	AAV35114	AAV35114 Mouse WRN
39	33.4	2.4	6476	2	AAH24302	AAH24302 Murine mW
40	33.4	2.4	7189	4	AAH30652	AAH30652 DNA encod
41	33.4	2.4	7189	4	AAH28725	AAH28725 Genomic s
42	33.4	2.4	7189	8	ACA03415	ACA03415 DNA encod
43	33.4	2.4	7189	9	ADB96763	ADB96763 Novel lun
44	33.4	2.4	7189	10	ADG41921	ADG41921 Human res
45	33.4	2.4	7189	11	ADI97695	ADI97695 Human res

ALIGNMENTS

RESULT 1	
AAf79834	
ID	AAf79834 strand; DNA; 1378 BP.
XX	AAf79834;
AC	
XX	
DT	30-MAY-2001 (first entry)
XX	
DE	Mycelia sterilia promoter sequence.
XX	
KW	Promoter; terminator; regulatory region; filamentous fungus;
KW	protein production; ds.
OS	Mycelia sterilia.
XX	
PN	WO200118219-A1.
XX	
PD	15-MAR-2001.
XX	
PF	07-SEP-2000; 2000WO-JP006104.
XX	
PR	07-SEP-1999; 990P-00252851.
XX	
PA	(MEIJU) MEIJU SEIKA KAISHA LTD.
XX	
PI	Watanabe M., Murakami T;
XX	
DR	WPI; 2001-235202/24.
XX	
PT	New promoter and terminator functioning synchronously, useful for
PT	regulating expression of endogenous gene in filamentous fungi,
PT	particularly transformed Mycelia sterilia capable of producing high
PT	yields of target protein/substance.
XX	
PS	Claim 2; Page 18-20; 26pp; Japanese.
XX	
CC	The present invention provides the sequences of a promoter and terminator
CC	from Mycelia sterilia which can be used to regulate the expression of an
CC	endogenous gene in filamentous fungi, particularly transformed Mycelia
CC	sterilia, to produce useful target proteins or other substances. The
CC	present sequence is the promoter of the invention
XX	
SQ	Sequence 1378 BP; 396 A; 289 C; 340 G; 353 T; 0 U; 0 Other;

Query Match	100.0%;	Score 1378;	DB 5;	Length 1378;
Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;
Matches 1378;	Conservative 0;	0;	0;	0;
QY	1	GTCCGCGGGGGGATATCATGCTGGTGTCCAAAACGCTGTTAGTGCATCGAT	60	
Db	1	GTCCGCGGGGGGATATCATGCTGGTGTCCAAAACGCTGTTAGTGCATCGAT	60	
QY	61	GAGGAAGAAACGCTGTGTGTGTGCGAGCTGAAGA CTGAAGAAGAGCCAAAGATATTC	120	
Db	61	GAGGAAGAAACGCTGTGTGTGTGCGAGCTGAAGA CTGAAGAAGAGCCAAAGATATTC	120	
QY	121	ACAATGCGATACGCTTGCATCATGCTTGTTCAGAGAGCAGTTGCATCTACCTGGTGT	180	
Db	121	ACAATGCGATACGCTTGCATCATGCTTGTTCAGAGAGCAGTTGCATCTACCTGGTGT	180	
QY	181	TCCCGCTTTCGTTGACAAAGATACAGTATCGGATGACCCGCCGCCGAAAGGAAATCCG	240	
Db	181	TCCCGCTTTCGTTGACAAAGATACAGTATCGGATGACCCGCCGCCGAAAGGAAATCCG	240	
QY	241	GAGTTCAAAAGAGGGTGTGCTCTACGCGCATTTAGTATAGATGCGATAGGGTTTGAAGTAA	300	
Db	241	GAGTTCAAAAGAGGGTGTGCTCTACGCGCATTTAGTATAGATGCGATAGGGTTTGAAGTAA	300	
QY	301	GCTGAAAGCTGATTAACGACATGACACAAAGAAATCAACGCTTGTATGCTTCCGCT	360	
Db	301	GCTGAAAGCTGATTAACGACATGACACAAAGAAATCAACGCTTGTATGCTTCCGCT	360	
QY	361	GCTTACTAAGATATCCAAAGACAAACAGCCGAAAGAAACCGATGCTGTAGAGGGG	420	
Db	361	GCTTACTAAGATATCCAAAGACAAACAGCCGAAAGAAACCGATGCTGTAGAGGGG	420	
QY	421	TTCCCTTTAGAGTCTACATGCTAAACGCTGATGATGAAGAAATCAATAGGCCAATCAAGTT	480	
Db	421	TTCCCTTTAGAGTCTACATGCTAAACGCTGATGATGAAGAAATCAATAGGCCAATCAAGTT	480	
QY	481	AGTATACCTGACGCTACATCGCTTTCTCCGATCTTGCCTTAAATATATGCTGCTGTC	540	
Db	481	AGTATACCTGACGCTACATCGCTTTCTCCGATCTTGCCTTAAATATATGCTGCTGTC	540	
QY	541	GAACTGCGGTACTGCTTCTGTAACGCTTCTCCGTTGAAGTCCCTAGGAGCAAGGCGCG	600	
Db	541	GAACTGCGGTACTGCTTCTGTAACGCTTCTCCGTTGAAGTCCCTAGGAGCAAGGCGCG	600	
QY	601	CGTTTGTAGACCTACATGATGCGACATCTTAAGAGAGGATCTGAGACATTTTCTAAGGC	660	
Db	601	CGTTTGTAGACCTACATGATGCGACATCTTAAGAGAGGATCTGAGACATTTTCTAAGGC	660	
QY	661	ATCCATATAGGCAATGGCGCTTAAGTCCGCAATTTGAAGAGATPAAGGGGGGTGTGAAGTG	720	
Db	661	ATCCATATAGGCAATGGCGCTTAAGTCCGCAATTTGAAGAGATPAAGGGGGGTGTGAAGTG	720	
QY	721	GTGTGTCAAAAGAGAGTGCATTTGCTATACAGCCGCTAAGAGAGGGGCTTGCAGCCTGT	780	
Db	721	GTGTGTCAAAAGAGAGTGCATTTGCTATACAGCCGCTAAGAGAGGGGCTTGCAGCCTGT	780	
QY	781	CTGCACTGTGTAATACGCTCACTTGTGTATGATGCCACTATATGTCCAGAGATGCAAA	840	
Db	781	CTGCACTGTGTAATACGCTCACTTGTGTATGATGCCACTATATGTCCAGAGATGCAAA	840	
QY	841	TGCTGATTTGGTTAAATATGCGCATGTATGATGATGTCGCAAAACAGTTTAAATCTAAGTT	900	
Db	841	TGCTGATTTGGTTAAATATGCGCATGTATGATGATGTCGCAAAACAGTTTAAATCTAAGTT	900	
QY	901	AAAGGAAGCTGAAGCTGAACCTGTCAGAAATTAAGCCCTGTGAATTAAGAGTTGATTA	960	
Db	901	AAAGGAAGCTGAAGCTGAACCTGTCAGAAATTAAGCCCTGTGAATTAAGAGTTGATTA	960	
QY	961	CCCAATTGAGTGTGTCAGAGGGTGTCTGATATGCTGAGAGCTTCCCTGTCCCATTTGGGGGT	1020	
Db	961	CCCAATTGAGTGTGTCAGAGGGTGTCTGATATGCTGAGAGCTTCCCTGTCCCATTTGGGGGT	1020	

QY	1021	AACTATTTGATAGTGGGCGAGAAATGCACCTGATTTTCAATGAACTTAACTATTCTGG	1080
Db	1021	AACTATTTGATAGTGGGCGAGAAATGCACCTGATTTTCAATGAACTTAACTATTCTGG	1080
QY	1081	GTAGAGATTTCTCAATAGGCTCTTCTGCTGTCACTTACACATCATAGGGGTCACAAAGT	1140
Db	1081	GTAGAGATTTCTCAATAGGCTCTTCTGCTGTCACTTACACATCATAGGGGTCACAAAGT	1140
QY	1141	ATACAGCTTTCATPAGAGTGGCGCATTTGAAATGATGCTACCGATCGAACCCGAAACGGTT	1200
Db	1141	ATACAGCTTTCATPAGAGTGGCGCATTTGAAATGATGCTACCGATCGAACCCGAAACGGTT	1200
QY	1201	CAAGCATGGGCGCTACGATGATCATATAGATGATGAAACATTAAGAGAGCTTGAAGAAC	1260
Db	1201	CAAGCATGGGCGCTACGATGATCATATAGATGATGAAACATTAAGAGAGCTTGAAGAAC	1260
QY	1261	CATTCAATTCCTAAGGGGCTCTCTTCTTGTGATCATCATCAAGAAATCATACCTCAAC	1320
Db	1261	CATTCAATTCCTAAGGGGCTCTCTTCTTGTGATCATCATCAAGAAATCATACCTCAAC	1320
QY	1321	CAGGAACCTTTTCTATCTTCCCTATPAGCAATTCGCAAAACCCATCAATCAACTTACA	1378
Db	1321	CAGGAACCTTTTCTATCTTCCCTATPAGCAATTCGCAAAACCCATCAATCAACTTACA	1378

RESULT 2
ADA71938
ID ADA71938 standard; DNA; 2000 BP.
XX
AC ADA71938;
XX
XX 20-NOV-2003 (first entry)
DT
DE Rice gene, SEQ ID 5263.
XX
XX Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
OS Oryza sativa.
XX
PN WO200300898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;
XX
DR MPI; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 27; SEQ ID NO 5263; 899bp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
XX involved in plant resistance or response to pathogenic infection. M1
XX comprises identifying a gene whose expression is significantly altered in
XX the incompatible interaction of plant gene expression relative to
XX expression of the gene in an uninfected plant, in a mutant plant that
XX does not express a gene associated with response to pathogenic infection,
XX or in a corresponding incompatible or compatible interaction. (M1) is
XX useful for conferring resistance to resistance or tolerance to a plant to
XX bacterial, fungal or viral infection. The present sequence was used to
XX illustrate the invention.

50	Sequence	2000 BP,	336 A;	265 C;	284 G;	363 T;	0 U;	752 Other;
	Query Match		3.5%;	Score 48.2;	DB 8;	Length 2000;		
	Best Local Similarity		10.0%;	Pred. No.	0.00034;			
	Matches	79;	Conservative 339;	Mismatches 372;	Indels 1;	Gap 1		
OY	CCTAGGACAAGGCCCGCCGTTTGTAGAACCCTACATGATGCACACTTTAAAGCAGGATCTG	644						
Db	28 SCARMSRSMKMSMKSRYRCSSCGKCCKMTTRRSKMYSASBSAGTGSWSSGSYSYG	87						
OY	AGACATTTTTTAAGGCATCATATATGGCACTTGTTGGGCGCTAAGTCGGCATTTGAAGGATATA	704						
Db	88 KGMKKRYRISKMRGRRRGRMRSRMRWRGYRRCARSGRMAAGSGRMMGKSRMSYWMM	147						
OY	GGGGGGTGAAAAGTGATGTGTCAAAGAAGGCTGCATTGGCTATATCACAGCCGCTAAGCAG	764						
Db	148 CYARCSCSCKRKSKSGSWGKTCTRGARAGSGWSGAKTKSSMSMKMMSSCGHSGCR	207						
OY	GTGGGCTTAGCAGCTGTCTGCACCTGTAATTAACGTCACTTGCTTAGATATGTCCACTPA	824						
Db	208 RSAYSRYYGTSRKTYTKMYTSASRCBRAMVMTTSSYNACSSYTWCBSKRSMWKMR	267						
OY	TGTACGACGATGCCAATATGCTGATTTGGGTTAAATGGCATGTAGTGTAGTCCCAAAAAC	884						
Db	268 KMRSRSYGVWSWSYKMMCTIAVKSYSRWCYTRYGGGWRAATRATWGGYNSRMAATK	327						
OY	ACGTTTAATCATGTTTAAAGGAACCTGAAGCTGAACCTGCTCAGAAATAAACCTGTTGG	944						
Db	328 KMYYRWGKXGMRGWAGRMRRMSNCBMSKALCYWRWRMRNTRRRRAKKSRTSRBK	387						
OY	AATACAACGTGTATAACCAATTCAGTCGTCAAGGGTGTCTGATATGCTGAGCTTCCC	1004						
Db	388 RKMCRKRRKYKMRGSRMRSCKBARMMRCSRGRAMKWGCRGCMTCBMKSYGMRRKSW	447						
OY	TGTGCCATTGTGGGGTACTATTTCATAGTGGGCGCAGATGCAACTATTTTCATITGA	1064						
Db	448 KRAMSKYKMSHWYWRKKKCSRTTWGKTRGAMGTWGRCHRYKRSQMKRCCRMRWR	507						
OY	ATCTAACTATTCGGGTAGAGTTCCTCAATGATCTTCGCGCTGCATTCACACATCA	1124						
Db	508 MYRMRWKRYMSARTTMRYCARAKTYSLSAARARCMYRSGKYVMAGMMKRYKMYMYM	567						
OY	TGGGGGTCAACAACGTATACAGCTTCA-TAGAGAGTGGCGCATTTGAAGTAGTACCGCAT	1184						
Db	568 MWMYRKYSKSGWYCMASYASCMKSARAAKMCBSHMASMSMSRSCRKCAASKR	627						
OY	CGAACCCCGAAGCGGTTCAAGACATAGGGCGTACGTAGATACATAGAGTATGAAAAATA	1244						
Db	628 SSARKYAMWGMTTSSRMSRMSLYCYWKKWSMSSTCYMTMYMSXYTAKGYSWRYRY	687						
OY	AAAGAGGCTTGAAGAACCATTCAATACCTAAGAGTCTCTCTTCTTCGATCACAATCA	1304						
Db	688 RAMCMYMRWYRYYRISYMTWTYMAVTSSRRMATGKYSGRYWTSMYKCKCSMKYRSMW	747						
OY	GAATCATACACTCAACACGAGAACCTTTTATCTTCCCTATATAGCAATTCAAAAACCA	1364						
Db	748 YYMSWMAAKTMGMWRRYATRRMMMYYRSMKYYTCTMGMYYMMWRITMRKRYMYKCT	807						
OY	TCAATCAACCT 1374							
Db	808 KTYWYWSATYW 818							
RESULT 3								
ABN80257/c								
ID	ABN80257	standard;	DNA;	6197	BP.			
XX	ABN80257;							
XX	AC							
XX	DT	15-JUL--2002	(first entry)					
XX	DE	Human chemically modified disease associated gene SEQ ID NO 274.						

XX	Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
KW	heart disease; epilepsy; histone deacetylation; muscular dystrophy;
KW	dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
KW	antidiabetic; cyostatic; anticonvulsant; ds.
XX	
OS	Homo sapiens.
OS	Synthetic.
PN	WO200200927-A2.
XX	
PD	03-JAN-2002.
XX	
PF	02-JUL-2001; 2001WO-EP007536.
XX	
PR	30-JUN-2000; 2000DE-01032529.
PR	01-SEP-2000; 2000DE-01043826.
XX	
PA	(EPFG-) EPIDENOMICS AG.
XX	
PI	Olek A, Piepenbrock C, Berlin K;
XX	
XX	WPI; 2002-130908/17.
XX	
PT	Novel nucleic acid useful for diagnosis and therapy of diseases
PT	associated with development genes such as diabetes, comprises a sequence
PT	of a segment of chemically pretreated DNA of genes associated with
PT	development.
XX	
XX	Claim 1; SEQ ID NO 274; 27bp; English.
XX	
CC	The invention relates to a nucleic acid (I) comprising a sequence at
CC	least 18 bases in length of a segment of chemically pretreated DNA (II)
CC	of genes associated with development selected from 87 genes listed in the
CC	specification such as ACCPM, ADFN, or APD1 and comprising one of 350
CC	sequences (AAN79984-ABN80333) or their complements. The invention is
CC	useful for the diagnosis or therapy of diseases associated with
CC	development genes, in particular disease related to homeobox containing
CC	genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes
CC	associated with congenital heart disease, epilepsy, diseases related to
CC	histone deacetylation, Currarino syndrome, diseases related with the
CC	development of the brain and limb girdle muscular dystrophy and dwarfism.
CC	Oligomers specific to each of the genes are useful for detecting the
CC	nucleic acid state of all CpG dinucleotides within the 350 sequences or
CC	(II) and their complementary sequences, as primer oligonucleotides for
CC	the amplification of the 350 sequences, (II) and/or their complements and
CC	as oligomer probes for detecting the cytosine methylation state and/or
CC	single nucleotide polymorphisms (SNPs). Note: The sequence data for this
CC	patent did not form part of the printed specification but is based on
CC	sequence information supplied to Derwent by the European Patent Office
XX	
XX	Sequence 6197 BP; 1637 A; 45 C; 1206 G; 3309 T; 0 U; 0 Other;
XX	
XX	Query Match 2.9%; Score 40; DB 6; Length 6197;
XX	Best Local Similarity 51.7%; Pred. No. 0.31; Mismatches 0; Gaps 0;
XX	Matches 91; Conservative 0; Indels 85; Indels 0; Gaps 0;
XX	
QY	1200 TCAGACATGGGCGGACGTAGATCATCATGAGTCAATGAAAGAGACTTGAAGA 1259
DB	3228 TAAATTAATTAACAACAACAACAACAACAACAACAACAACAACAATTAATTAACAA 3159
QY	1260 CCATTCAATCTTAAGGGCTCTCTCTTTCTTGATCATCATCAAGATCATACACTCAA 1319
DB	3168 CAATTAATAATCTTAATCTTCTCTTTAATAAAAAAATTCTTAATAAACAATACCT 3109
QY	1320 CCAGAACTCTTTCTATCTTCCCTATAGCAATTCGCAAAACCCTCATCACTA 1375
DB	3108 CAATAAACTCATTTTAATTAATTAACCTCCCTTTTACACAAAAAAGTACCAATCTA 3053
XX	
XX	RESULT 4
XX	ADA71938/c
XX	ID ADA71938 standard; DNA; 2000 BP.

AC		ADA71938;
PC		
DT	XX	20-NOV-2003 (first entry)
DE	XX	Rice gene, SEQ ID 5263.
KM	XX	Plant; bacterial infection; fungal infection; viral infection; rice; gene; ds.
OS	XX	Oryza sativa.
PN	XX	WO2003000898-A1.
PD	XX	03-JAN-2003.
PF	XX	22-JUN-2001; 2001WO-IB001105.
PR	XX	22-JUN-2001; 2001WO-IB001105.
PA	XX	(SYGN) SYNGENTA PARTICIPATIONS AG.
PI	XX	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y, Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G; WPI; 2003-175290/17.
PT	XX	Identifying at least one gene involved in plant resistance or response to pathogenic infection for conferring resistance or tolerance to a plant to bacterial, fungal or viral infection by determining or detecting plant gene expression.
PS	XX	Claim 27; SEQ ID NO 5263; 899pp; English.
CC	XX	The present invention relates to a method (M1) for identifying genes involved in plant resistance or response to pathogenic infection. M1 comprises identifying a gene whose expression is significantly altered in the incompatible interaction of plant gene expression relative to expression of the gene in an uninfected plant, in a mutant plant that does not express a gene associated with response to pathogenic infection, or in a corresponding incompatible or compatible interaction. (M1) is useful for conferring resistance to resistance or tolerance to a plant to bacterial, fungal or viral infection. The present sequence was used to illustrate the invention.
SQ	XX	Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;
		Query Match 2.7%; Score 37.2; DB 8; Length 2000; Best Local Similarity 10.2%; Pred. No. 1.4; Matches 2; Gaps 1 Matches 42; Conservative 190; Mismatches 178; Indels 2; Gaps 1
DG	XX	362 CTCTAAGATGATAATCCAAGAGACAACAGCGGAAGAACCAGTGTCTGAGGGGT 421
DB	XX	:: :: :::: : : : : : : : : : : : : : : : : : : : : : : : : : : 557 MKMKCTWRRCMCYRNGCYMTYTTSRSMMYTGTGYKARYTSKRHYMYKKRYCKMYYYGNY 498
DG	XX	422 TCCTTTAGAGTGTCATGTGTAACGCGTCAATGATAGAAATCATCAATGCCAATCAAAGTTA 481
DB	XX	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : 497 MKCSYMRRGYCKACKKCXYAMC--WKAAYSGMWWYWVKYSKMWMSTKYMWSMWYKXC 440
DG	XX	482 GTAATACGCCTACTCATGCTCTTTCTCCGGAATCTTGCTTAATAATATATGCTGCTCCG 541
DB	XX	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : 439 RSMRKGA KCAYCCCKMYTCSYGEMKMYTYMGSKYSRCKYKRMWYMVMKMMYMSAYS 380
DG	XX	542 AACGTGCGGTACTGCTCTGCTACTAACGTGTTCTTCGTTAAGGCTTAGAGACAAGCGCGG 601
DB	XX	:: :: :::: : : : : : : : : : : : : : : : : : : : : : : : : : : 379 SMMTWIYYIAKWTKWIKYRRGMTSWYGSXKYYLCTWCYCMKRCYRRKMKRKTXTKS 320
DG	XX	602 GTTTGTAAGCTPACATGATGACACATCTTAAAGCAGGAGTCTGAGACATTTTCTAAGCA 661
DB	XX	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : 319 KRCYCGRVATCTWC CCYRKRWYSRRSMMRTGAKMWRBSWSMBRSYSWYKMYKKMWKKS 260
DG	XX	662 TTCATATAGGCAATTGGCGCCTAAGTCGCAATTTGAAGAGATTAAGGGGGGTGAAAAGTG 721
DB	XX	:: :: :::: : : : : : : : : : : : : : : : : : : : : : : : : : : 259 YTMYSGVMASSSGTWSAAKRTYTKYSTBRRAKMMBACMYSA CRBYSTRSTSYCGCYSG 200

[illegible]

Dh 642 MYSMGKTITTCRAWMKSSGYTKCGTGGTAGYWCAMWWSCRMWMSCRGSGKSCCTTKWT 583
Qy 794 TAAGCTCACTTGCTTAGGTATGTCCACCTAATGTACAGACGATGCAAAATGC 843
Dh 582 KAAMRAMMCTYYWTGSGAAMAASYTKYKAGGKMKCGKGCRCRAMTKY 533

RESULT 6

ADN96821/c
ID ADN96821 standard; DNA, 2672 BP.

AC ADN96821;

DT 01-JUL-2004 (first entry)

DE Bugula neritina PKS cosmid clone 5B Pat D4/C1 overlap DNA SegID 34.

XX ds; degenerate; polyketide; bryopyran ring; bryostatin;

KW polyketide synthase; PKS; toxic; cytostatic; immunomodulatory;

KM protein therapy; cancer; metastasis.

XX Bugula neritina.

OS WO200309219-A2.

XX PD 04-DEC-2003.

XX PF 20-MAY-2003; 2003WO-US016299.

XX PR 20-MAY-2002; 2002US-0382181P.

XX PA (REGC) UNIV CALIFORNIA.

PA (HAYG/) HAYGOOD M.

PA (HILD/) HILDEBRAND M.

PA (ANDE/) ANDERSON C.

PA (WAGG/) WAGGONER L E.

PA (SHER/) SHERMAN D H.

PA (LITH/) LITH H.

XX Haygood M, Hildebrand M, Anderson C, Waggoner LE, Sherman DH;

PI Liu H;

XX WPI; 2004-053143/05.

XX DR

XX New compositions comprising a polyketide synthase or nucleic acid

FT encoding the polyketide synthase, useful in biosynthesizing polyketides,

FT bryopyran rings and bryostatins having anti-cancer or antimetastatic

FT activity.

XX Claim 13; SEQ ID NO 34; 342bp; English.

XX This invention relates to a novel composition that comprises at least one

CC polypeptide that catalyzes the one step synthesis of a polyketide or

CC bryopyran ring. Specifically, it refers to nucleic acid molecules derived

CC from marine organisms that encode enzymes that catalyze the synthesis of

CC bioactive compounds such as polyketides and bryostatins that are based on

CC the use of polyketide synthases (PKSs) to generate toxic polyketides that

CC exhibit cytostatic and immunomodulatory activities, such that they can be

CC used for protein therapy in the treatment of cancer and metastasis. This

CC polynucleotide is a Bugula neritina PKS DNA sequence of the invention.

XX Sequence 2672 BP; 686 A; 502 C; 570 G; 660 T; 0 U; 254 Other;

Qy Query Match 2.6%; Score 36.2; DB 12; Length 2672;

Best Local Similarity 24.1%; Pred. No. 3.4;

Matches 41; Conservative 60; Mismatches 69; Indels 0; Gaps 0;

Qy 674 TTGGCGCTAGTCGACATTGAAGAGATTAAGGCGGCTGTGAAGTGTCTCAAAAG 733

Dh 702 TTCWGMGCMRSTCAKSMYKTKWTWCCCAAGSTGRKSKYKOTWTKWTYSGWMANCRG 643

Qy 734 AGTCGATTCGCTATACAGCCGCTAAGCAGGTGGCTAGACGTGTCTGCAGCTGTGAA 793
Dh 642 MYSMGKTITTCRAWMKSSGYTKCGTGGTAGYWCAMWWSCRMWMSCRGSGKSCCTTKWT 583
Qy 794 TAAGCTCACTTGCTTAGGTATGTCCACCTAATGTACAGACGATGCAAAATGC 843
Dh 582 KAAMRAMMCTYYWTGSGAAMAASYTKYKAGGKMKCGKGCRCRAMTKY 533

RESULT 7

AAH24065
ID AAH24065 standard; DNA, 4590 BP.

AC AAH24065;

DT 29-AUG-2001 (first entry)

DE Yeast AOD9604-associated DNA sequence, SEQ ID NO:1.

XX Human growth hormone analogue peptide; hGH; AOD9604; lipid metabolism;

KW modulation; lipolysis stimulation; hormone-sensitive lipase stimulation;

KM lipogenesis inhibition; acetyl CoA carboxylase inhibition; obesity;

KW functional food; transgenic yeast; fat/lean ratio; food use; ds.

OS Saccharomyces cerevisiae.

XX Key Location/Qualifiers

FT misc_feature 10

FT /tag= a

FT /note= "Represented as * in the specification"

FT misc_feature 3617

FT /tag= b

FT /note= "Represented as * in the specification"

FT misc_feature 3649

FT /tag= c

FT /note= "Represented as * in the specification"

FT misc_feature 3679

FT /tag= d

FT /note= "Represented as * in the specification"

FT misc_feature 3819

FT /tag= e

FT /note= "Represented as * in the specification"

FT misc_feature 3862

FT /tag= f

FT /note= "Represented as * in the specification"

FT misc_feature 3864

FT /tag= g

FT /note= "Represented as * in the specification"

FT misc_feature 3888

FT /tag= h

FT /note= "Represented as * in the specification"

FT misc_feature 3890

FT /tag= i

FT /note= "Represented as * in the specification"

FT misc_feature 3912

FT /tag= j

FT /note= "Represented as * in the specification"

FT misc_feature 3914

FT /tag= k

FT /note= "Represented as * in the specification"

FT misc_feature 3938

FT /tag= l

FT /note= "Represented as * in the specification"

FT misc_feature 3939

FT /tag= m

FT /note= "Represented as * in the specification"

FT misc_feature 3941

FT /tag= n

FT /note= "Represented as * in the specification"

FT misc_feature 3943

FT /tag= o

FT /note= "Represented as * in the specification"

FT misc_feature 4361

FT /tag= p

FT /note= "Represented as * in the specification"

QY 482 GTATACCTGACGCTACATC 500
Db 118373 GTACAGCTGACCCCTTCAC 118355

RESULT 9

AAF83969/c
ID AAF83969 standard; DNA, 993 BP.

AC AAF83969;

DT 06-AUG-2001 (first entry)

XX Human SER4 nucleotide sequence.

XX Serine/threonine kinase-like protein; SERX; breast; testicular; cancer;
KM cardiovascular; autoimmune disorder; cytostatic; cardiac;
KM gene therapy. immunosuppressive; antiaesthetic; antifibrinolytic; SER4;
de.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..993

FT /*tag= a

FT /product= "SER4"

XX MO200136645-A2.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000MO-US031744.

XX 17-NOV-1999; 99US-0165986P.

PR 09-FEB-2000; 2000US-0181347P.

PR 03-APR-2000; 2000US-0194195P.

PN 05-APR-2000; 2000US-0194839P.

PR 07-APR-2000; 2000US-0195637P.

PR 13-APR-2000; 2000US-0197080P.

PR 15-SEP-2000; 2000US-0232677P.

PR 16-NOV-2000; 2000US-00715427.

XX (CURA-) CURAGEN CORP.

XX Quinn KE, Szytek KA, Majumder K, Vernet C, Burgess C;

PI Fernandes E, Taupier R, Rastelli L, Herrmann JL;

XX WPI; 2001-336006/35.

DR P-PSDB; AAB85038.

XX New isolated SERX polynucleotides and polypeptides related to the

PT serine/threonine kinase family of proteins, useful for determining and
PT treating diseases related to altered levels of the polypeptide.

XX Claim 9; Page 21; 127pp; English.

XX The invention relates to serine/threonine kinase-like proteins, termed
CC SERX. The SERX polynucleotides can be used to express SERX protein (e.g.

CC applications), to detect SERX mRNA (e.g. in a biological sample) or a
CC genetic lesion in a SERX gene and modulate SERX activity. The SERX
CC polypeptides, polynucleotides, antibodies can also be used to treat or
CC prevent a pathology associated with SERX. Disorders associated with
CC aberrant SERX expression or activity include breast or testicular cancer,
CC cardiovascular defect, autoimmune disorders and asthma. Disorders
CC associated with aberrant SERX-8 expression include, blood clotting
CC hematopoietic and tumor related disorders. The present sequence
CC represents the nucleotide sequence of human SER4

XX Sequence 993 BP; 229 A; 305 C; 264 G; 195 T; 0 U; 0 Other;

Query Match 2.6%; Score 35.4; DB 4; Length 993;

Best Local Similarity 54.1%; Pred. No. 3.6;

Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 904 GGAAGCTGAAAGCTGAACCTGTGAAATGACCTGTGGAATCAACGTTGATTAACCC 963

Db 708 GGAAGCTCAAGGCTGTACCGGTCAATGAGAGGCTGCGGCGAGGCAATGATTTCTC 649

QY 964 AATTCAGTCGTCAAGGCTGTCTGTATGCTGAGCTTCCCTGTGCAATTGGGGTAA 1023

Db 648 AGTTGTCCCTGAGAGGTGATACATCACCGCAGCCTCGAAAGCCCTTGAAGCTGGC 589

QY 1024 TATTTCAATGAG 1036

Db 588 CAGCTCGTAGAG 576

RESULT 10

ADE29362/c
ID ADE29362 standard; cDNA, 993 BP.

AC ADE29362;

DT 29-JAN-2004 (first entry)

XX Human cDNA encoding serine/threonine kinase SER4.

XX Human; ss; gene; serine threonine kinase; SER4;

KM blood coagulation disorder; liver disorder; hepatoma cell disorder;

KM hepatocellular cell disorder; obesity; diabetes; cancer; gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..993

FT /*tag= a

FT /product= "SER4"

XX US2003077697-A1.

XX 24-APR-2003.

XX 03-JUL-2001; 2001US-00898837.

XX 17-NOV-1999; 99US-0165986P.

PR 09-FEB-2000; 2000US-0181347P.

PR 03-APR-2000; 2000US-0194195P.

PR 07-APR-2000; 2000US-0194839P.

PR 13-APR-2000; 2000US-0195637P.

PR 03-JUL-2000; 2000US-0215906P.

PR 15-SEP-2000; 2000US-0232677P.

XX (GERL/) GERLACH V L.

PA (MACD/) MACDOUGALL J R.

PA (QUIN/) QUINN K E.

PA (MAJU/) MAJUMDER K.

PA (SPYT/) SPYTEK K A.

PA (VERN/) VERNER C.

PA (BURG/) BURGESS C E.

PA (FERN/) FERNANDES E R.

PA (RAST/) RASTELLI L.

PA (HERR/) HERRMANN J L.

PA (SPAD/) SPADERNA S K.

PA (SHIM/) SHIMKETS R A.

PA (TAUP/) TAUPIER R J.

XX Gerlach VV, MacDougall JR, Quinn KE, Majumder K, Szytek KA;

PI Vernet C, Burgess CE, Fernandes ER, Rastelli L, Herrmann JL;

PI Spaderna SK, Shimkets RA, Taupier RJ;

XX WPI; 2003-615999/58.

DR P-PSDB; ADE29363.

XX New serine/threonine protein-kinase like polypeptides and genes, useful

PT in gene therapy, diagnosis or prognosis of e.g. diseases of blood
PT coagulation, obesity; diabetes, or cancers.
PS Claim 9; SEQ ID NO 8; 77pp; English.
XX
CC The invention relates to a new isolated polypeptide has a sequence
CC (designated SER1, SER2, SER4, SER5 and SER6, respectively) fully defined
CC in the specification (or their mature forms, variants or fragments). Also
CC included are an isolated nucleic acid molecule (comprising: a nucleic
CC acid sequence encoding any of the SER polypeptides above; a nucleic acid
CC fragment encoding at least a portion of the SER polypeptides or their
CC variants; or their complements), vectors comprising the nucleic acid
CC molecules, cells comprising the vectors, an antibody that binds
CC immunospecifically to the SER polypeptide, determining the presence or
CC amount of the SER polypeptide or nucleic acid in a sample, identifying an
CC agent that binds to the SER polypeptide, identifying a (potential)
CC therapeutic agent for use in the treatment of a pathology related to
CC aberrant expression or aberrant physiological interactions of the SER
CC polypeptide, modulating the activity of the SER polypeptide,
CC pharmaceutical compositions comprising a carrier (and the SER
CC polypeptide, nucleic acid molecule or antibody), kits comprising in one
CC or more containers any of the compositions above, screening for a
CC modulator of activity or latency or predisposition to a pathology
CC associated with the SER polypeptide and determining the presence of or
CC predisposition to a disease associated with altered levels of the SER
CC polypeptide or nucleic acid molecule. The SER polypeptide, nucleic acid
CC or anti-SER antibody is useful for treating or preventing a pathological
CC state or pathology associated with the SER polypeptide in a subject,
CC particularly a human. These pathologies include diseases of blood
CC coagulation, human liver, hepatoma cells or hepatocellular cells,
CC obesity, diabetes, or cancers. The polypeptide or polynucleotide is also
CC useful in gene therapy, diagnostics, prognosis, or research. The present
CC sequence encodes SER4.
XX
SQ Sequence 993 BP; 229 A; 305 C; 264 G; 195 T; 0 U; 0 Other;
Query Match 2.6%; Score 35.4; DB 10; Length 993;
Best Local Similarity 54.1%; Pred. No. 3.6;
Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 904 GGAAGCGTGAAGCTGTCAGAAATGAGCTGTGGAATPACAGTTGATTAACCC 963
DB 708 GCGAAGCTCAGAGGCTGATCCGTCAGATPAGAGGCTGCGCGAGGAGATTTTTC 649
QY 964 AATTCACTGTCAGAGGCTGCTGATATGCTGAGCTTCCTGTGCAATGCGGGTAAC 1023
DB 648 AGTTGTCCCTCGAGAGGTCATACACACCGCAGACCTCGAAAGCCCTTGAGGCTGGC 589
QY 1024 TATTTCATATGCG 1036
DB 588 CAGCTCGTACAGAG 576
RESULT 11
ABL3237/c
ID ABL3237 standard; DNA; 6971 BP.
XX
AC ABL32327;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 1210.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytoskeletal; noctropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antineumatic; antirheumatic; antidiabetic; antiporiatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
de.
XX
XX Homo sapiens.
OS

XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP007537.
PF 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PT Olek A, Piepenbrock C, Berlin K;
DR WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
PS
XX Claim 1; SEQ ID NO 1210; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX
SQ Sequence 6971 BP; 1665 A; 242 C; 1831 G; 3232 T; 0 U; 1 Other;
Query Match 2.5%; Score 35; DB 6; Length 6971;
Best Local Similarity 53.2%; Pred. No. 14;
Matches 74; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
QY 1238 AACATAAAGAGGCTTGAAGAACCATTCATTAAGGCTCTTTCTTTCGATCA 1237
DB 2068 AAAAAAAAAAACTATATAAAACCTAAAAACCTCCAGACCTCTTACTTCTTAACA 2009
QY 1298 CATCAAGATCATCTACTCAAAACGAGAACTTTTATCTTCGCTATAGCAATTTCCAA 1357
DB 2008 AACTATATPACTATACAAAAATCCAAAAAAATCTTAACTCCCTCACTCCACCCCA 1949
QY 1358 AACCATCAATCAACCTAA 1376
DB 1948 CAACATCCCCCAACTAAA 1930
RESULT 12
AB272040/c
ID AB272040 standard; DNA; 207433 BP.
XX
AC AB272040;
XX
DT 03-APR-2003 (first entry)
XX
DE Gene 216 H194BAC1098L22 nucleotide sequence SEQ ID NO 5.
XX
KW Human; Gene 216; chromosome 20p13-p12; antiasthmatic; anorectic;
KW antiinflammatory; gastrointestinal; gene therapy; vaccine; asthma;
KW obesity; inflammatory bowel disease; promoter; gene; ss.
XX
OS Homo sapiens.
XX
XX WO200178894-A2.
FN 25-OCT-2001.
XX
PD 13-APR-2001; 2001WO-US012245.
XX
PF 13-APR-2000; 2000US-00548797.
PR

XX	(GENO-) GENOME THERAPEUTICS CORP.
PA	
XX	
P1	Kelith T,
DR	WPI; 2001-639428/73.
DR	P-PsDB; ABR00926.
XX	
PT	Isolated genes (Gene 216) from human chromosome 20p13-p12 and the
PT	proteins they encode, useful for the prevention, diagnosis and treatment
XX	of asthma, obesity and inflammatory bowel disease.
PS	
XX	Example 4; Fig 7; 520pp; English.
CC	The invention relates to isolated genes (Gene 216) from human chromosome
CC	20p13-p12 and the proteins they encode. The nucleic acids and proteins
CC	may be used in the prevention, diagnosis and treatment of diseases
CC	associated with inappropriate Gene 216 expression. For example, the
CC	nucleic acids (or vectors) and proteins may be used to treat disorders
CC	associated with decreased expression by rectifying mutations or deletions
CC	in a patient's genome that affect the activity of gene 216 by expressing
CC	inactive proteins or to supplement the patients own production of Gene
CC	216 proteins. Additionally, the nucleic acids may be used to produce the
CC	secreted Gene 216 protein, by inserting the nucleic acids into a host
CC	cell and culturing the cell to express the protein. The nucleic acids and
CC	complementary sequences may also be used as DNA probes in diagnostic
CC	assays to detect and quantitate the presence of similar nucleic acid
CC	sequences in samples and therefore which patients may be in need of
CC	restorative therapy. The Gene 216 protein may also be used as antigens in
CC	the production of antibodies against Gene 216 and in assays to identify
CC	modulators of Gene 216 expression and activity. The anti-gene 216
CC	antibodies and antagonists may also be used to down regulate expression
CC	and activity. The anti-gene 216 antibodies may also be used as diagnostic
CC	agents for detecting the presence of Gene 216 proteins in samples (e.g.
CC	by enzyme linked immunosorbent assay or ELISA). Disorders that may be
CC	prevented, diagnosed and/or treated by the above methods include, for
CC	example asthma, obesity and inflammatory bowel disease. The present
CC	sequence is that of the Gene 216 genomic nucleic acid sequence, promoter
CC	or enhancer
XX	
SQ	Sequence 207433 BP; 52775 A; 51290 C; 51698 G; 51670 T; 0 U; 0 Other;
XX	
Query Match	2.5%; Score 35; DB 5; Length 207433;
Best Local Similarity	55.3%; Pred. No. 97;
Matches	68; Conservative 0; Mismatches 55; Indels 0; Gaps 0
QY	1210 GGCGTACGTAGTACTAGATGTCATRGAAACATAAAGAAGCTTGAAACAACATTCAAT 1265
DB	1335 GTCTTACCCTCTTTCTTAACACAAATACACMAATTCACAGGGCTAGTAGAACCTTGAAAT 1270
QY	1270 CCTAAGGGCTCCTCTTCTTCTTGCAATCACATCAAGAACTCATCACTCAAAACGAAGATC 1325
DB	1275 ATACGAACATCTCTTCTTCAAAGAAAACCTCAGGATATTAATCCAATRGAAACCC 1210
QY	1330 TTT 1332
DB	1215 ATT 1213
RESULT 13	
ID	ABX74891/C
ID	ABX74891 standard; DNA; 207433 BP.
AC	ABX74891;
DT	07-APR-2003 (First entry)
DE	BAC1098L22 DNA sequence.
XX	
KM	Gene 216; antiasthmatic; antiinflammatory; ss; anorectic;
KM	chromosome 20p13-p12; single nucleotide polymorphism; SNP; gene therapy;
KM	respiratory disease; asthma; obesity; bronchial hyper-responsiveness;
KM	chronic obstructive pulmonary disease;

KW		adult respiratory distress syndrome; inflammatory bowel syndrome.
XX		
OS		Synthetic.
XX		
PN		WO200283077-A2.
XX		
PD		24-OCT-2002.
XX		
PF		15-APR-2002; 2002WO-USO12063.
XX		
PR		13-APR-2001; 2001US-00834597.
XX		
PR		13-APR-2001; 2001WO-USO12245.
XX		
PA	(SCHE) SCHERING CORP.	
XX	(GENO-) GENOME THERAPEUTICS CORP.	
PI	Keith T, Little RD, Van Berdegh P, Dupuis J, Del Mastro RG;	
XX	Simon J, Allen K, Pandic S;	
XX	WPI; 2003-092960/08.	
XX		
PT	New isolated gene 216 nucleic acids, useful for diagnosing, preventing or	
XX	treating a disorder, such as asthma, bronchial hyper-responsiveness,	
XX	chronic obstructive pulmonary disease, obesity or inflammatory bowel	
XX	syndrome.	
PS	Example 6; Fig 7; 650bp; English.	
CC	This invention relates to a novel isolated nucleic acid, gene 216,	
CC	identified from human chromosome 20p13-p12. The invention also discloses	
CC	regions of the 216 gene that contain single nucleotide polymorphisms	
CC	(SNP's) which may be used as markers for disease susceptibility or	
CC	severity. The nucleotides of the invention may have antiasthmatic,	
CC	antiinflammatory or anorectic activities and may be used in gene therapy.	
CC	The nucleic acids, antibodies or its fragments are useful for diagnosing,	
CC	preventing or treating a disorder, such as respiratory diseases (e.g.	
CC	asthma, bronchial hyper-responsiveness, chronic obstructive pulmonary	
CC	disease or adult respiratory distress syndrome), obesity, or inflammatory	
CC	bowel syndrome. The nucleic acids are also useful for identifying	
CC	increased susceptibility of a subject to the disorders mentioned. The	
CC	nucleic acids can also be used as primers and templates for the	
CC	recombinant production of disorder-associated peptides or polypeptides,	
CC	for chromosome and gene mapping, or for tissue distribution studies. The	
CC	present sequence represents a gene 216 cDNA sequence used in the scope of	
CC	the invention	
XX		
SO	Sequence 207433 BP; 52775 A; 51289 C; 51698 G; 51671 T; 0 U; 0 Other;	
	Query Match	2.5%; Score 35; DB 8; Length 207433;
	Best Local Similarity	55.3%; Pred.No. 92;
	Matches	68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
OY	1210 GGCGTAGCATGATCATGAGTCTATGAAAACATTAAGAGAGCTGGAAACCATTCAAT	1265
DB	1335 GTCTTACTTTCTTTAAMACAAATACCAATTCGCCAGGGCTGTGAGACCTTGAAAT	1276
OY	1270 CCTAAGGTCCTCTCTTCTTGATGCATCACAAGAATCATACACTCAAACAGAACTC	1329
DB	1275 ATACGAACATCTCTTCTTAAGGAAACCTCAGGATATTAATAATCCAATAGAAACC	1216
OY	1330 TTT 1332	
DB	1215 ATT 1213	
RESULT 14		
ID	ADJ3614/C	
XX	ADJ3614 standard; DNA; 207433 BP.	
AC	ADJ3614;	
DT	22-APR-2004 (first entry)	
XX		

DE Bacterial artificial chromosome RBC1-11.
 XX antiasthmatic; respiratory; gene therapy; asthma;
 KW bronchial hyperresponsiveness; atopy; chronic obstructive lung disease;
 KW adult respiratory distress syndrome; obesity; inflammatory bowel disease;
 KW human; gene 216; bacterial artificial chromosome; BAC; RBC1-11; ds.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FN US2004002470-A1.
 XX
 PD 01-JAN-2004.
 XX
 PD 17-OCT-2002; 2002US-00277216.
 PF
 XX 13-APR-2000; 2000US-00548797.
 PR 13-APR-2001; 2001US-00834597.
 PR 19-APR-2002; 2002US-00126022.
 XX
 PA (KEIT/) KEITH T.
 PA (LITT/) LITTLE R D.
 PA (VEER/) VAN EERDEMEGH P.
 PA (DUPU/) DUPUIS J.
 PA (DMAS/) DEL MASTRO R G.
 PA (SIMO/) SIMON J.
 PA (ALLE/) ALLEN K.
 PA (PAND/) PANDIT S.
 XX
 PI Keith T, Little RD, Berdewegh PV, Dupuis J, Del Mastro RG;
 PI Simon J, Allen K, Pandit S;
 XX
 DR WPI; 2004-061675/06.
 XX
 PT Gene 216 nucleic acid, useful for preparing a composition for treating
 PT disorders e.g., asthma, bronchial hyperresponsiveness, atopy, chronic
 PT obstructive lung disease and adult respiratory distress syndrome.
 XX
 XX Example 6; SEQ ID NO 5; 441bp; English.
 XX
 CC The invention describes a new isolated nucleic acid comprising a fully
 CC defined sequence having 23574 bp or at least its 50 or 15 contiguous
 CC nucleotides and includes: allele G of single nucleotide polymorphism
 CC (SNP) AB+2; allele G of SNP BC+1; and allele C of SNP BC+2. The invention
 CC describes identifying increased susceptibility to a disorder comprising
 CC asthma, bronchial hyperresponsiveness, atopy, chronic obstructive lung
 CC disease and adult respiratory distress syndrome in a subject comprising
 CC testing a biological sample obtained from a subject for the presence of
 CC at least one allele or haplotype given in the specification, where the
 CC presence identifies an increased susceptibility to the disorder. The
 CC nucleic acid is useful for preparing a composition for treating disorders
 CC comprising asthma, bronchial hyperresponsiveness, atopy, chronic
 CC obstructive lung disease and adult respiratory distress syndrome. This
 CC sequence represents a bacterial artificial chromosome (BAC) containing
 CC human gene 216 isolated from chromosome 20p13-p12 associated with asthma,
 CC obesity and inflammatory bowel disease.
 CC
 XX
 SQ Sequence 207433 BP; 52775 A; 51290 C; 51698 G; 51670 T; 0 U; 0 Other;
 Query Match 2.5%; Score 35; DB 12; Length 207433;
 Best Local Similarity 55.3%; Pred. No. 92;
 Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
 QY 1210 GCCGTAAGTACATAGATGATGAAACATTAAGAGCTTGAAGAACATTCAAT 1269
 DB 1335 GTCTTACCTCTTCTTAACACAAATATCCAGGCTAGTAGAGACCTTGAAT 1276
 QY 1270 CCTAAGGCTCTCTTCTTCTGATCATCATCAAGATCATCACTCAACAGGAATCTC 1329
 DB 1275 ATACGAACATCTCTTCTTAAGAAACCTCAGGATATTAATTCACAAATAGAAACC 1216
 QY 1330 TTT 1332
 ||

DB 1215 ATT 1213
 RESULT 15
 ADL81193/c
 ID ADL81193 standard; DNA; 207433 BP.
 XX
 AC ADL81193;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE BAC1098L22 DNA sequence.
 XX
 KW asthma; bronchial hyperresponsiveness; obesity;
 KW inflammatory bowel disease; human; ds; gene 216.
 XX
 OS Homo sapiens.
 OS
 XX
 FN US2004023215-A1.
 XX
 PD 05-FEB-2004.
 XX
 PD 19-APR-2002; 2002US-00126022.
 PF
 XX 13-APR-1999; 99US-0129391P.
 PR 13-APR-2000; 2000US-00548797.
 PR 13-APR-2001; 2001US-00834597.
 XX
 PA (KEIT/) KEITH T.
 PA (LITT/) LITTLE R D.
 PA (VEER/) BERDEMEGH P V.
 PA (DUPU/) DUPUIS J.
 PA (DMAS/) DEL MASTRO R G.
 PA (SIMO/) SIMON J.
 PA (ALLE/) ALLEN K.
 PA (PAND/) PANDIT S.
 XX
 PI Keith T, Little RD, Berdewegh PV, Dupuis J, Del Mastro RG;
 PI Simon J, Allen K, Pandit S;
 XX
 DR WPI; 2004-142647/14.
 XX
 PT New isolated nucleic acid molecules useful for diagnosing or treating
 PT asthma or bronchial hyperresponsiveness, or other diseases such as
 PT obesity or inflammatory bowel disease.
 XX
 XX Example 5; SEQ ID NO 5; 485bp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule, or a set of
 CC nucleic acid molecules each given in the specification. The composition
 CC and methods are useful in diagnosing or treating asthma or bronchial
 CC hyperresponsiveness, and other diseases such as obesity or inflammatory
 CC bowel disease. The present sequence is used in the exemplification of the
 CC present invention.
 CC
 XX
 SQ Sequence 207433 BP; 52775 A; 51290 C; 51698 G; 51670 T; 0 U; 0 Other;
 Query Match 2.5%; Score 35; DB 12; Length 207433;
 Best Local Similarity 55.3%; Pred. No. 92;
 Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
 QY 1210 GCCGTAAGTACATAGATGATGAAACATTAAGAGCTTGAAGAACATTCAAT 1269
 DB 1335 GTCTTACCTCTTCTTAACACAAATATCCAGGCTAGTAGAGACCTTGAAT 1276
 QY 1270 CCTAAGGCTCTCTTCTTCTGATCATCATCAAGATCATCACTCAACAGGAATCTC 1329
 DB 1275 ATACGAACATCTCTTCTTAAGAAACCTCAGGATATTAATTCACAAATAGAAACC 1216
 QY 1330 TTT 1332
 DB 1215 ATT 1213

, Fri Feb 25 18:19:08 2005

us-10-070-386-1.rng

Page 11

Search completed: February 9, 2005, 10:48:33
Job time : 757 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 10:24:55 ; Search time 272 Seconds
(without alignment)
8289.667 Million cell updates/sec

Title: US-10-070-386-1
Perfect score: 1378
Sequence: 1 gtcacgctggctgctgatat.....accatcatcaactacaa 1378

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUG_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	42.2	3.1	7218	1 US-08-232-463-14	Sequence 14, Appl
C 2	37.6	2.7	289	3 US-09-007-005-17	Sequence 17, Appl
C 3	37.6	2.7	289	3 US-09-244-796-17	Sequence 17, Appl
C 4	35.2	2.6	399	4 US-09-621-976-8976	Sequence 8976, Ap
C 5	35.2	2.6	505	4 US-09-621-976-15639	Sequence 15639, A
C 6	35.2	2.5	136917	4 US-09-949-016-16369	Sequence 16369, A
C 7	34.4	2.5	832	4 US-09-621-976-2813	Sequence 2813, Ap
C 8	34.4	2.5	16871	4 US-09-949-016-12718	Sequence 12718, A
C 9	34.4	2.5	16874	4 US-09-949-016-15176	Sequence 15176, A
C 10	34.2	2.5	1500	4 US-09-107-532A-1870	Sequence 1870, Ap
C 11	34.2	2.5	134890	4 US-09-949-016-15602	Sequence 15602, A
C 12	33.4	2.4	4792	3 US-08-781-891-205	Sequence 205, App
C 13	33.4	2.4	4792	3 US-09-618-166-205	Sequence 205, App
C 14	33.4	2.4	6476	3 US-09-127-670-5	Sequence 5, Appl1
C 15	33.4	2.4	16442	3 US-08-781-891-208	Sequence 208, App
C 16	33.4	2.4	16442	4 US-09-618-166-308	Sequence 208, App
C 17	33.3	2.4	364	4 US-09-621-976-17202	Sequence 17202, A
C 18	32.8	2.4	414	4 US-09-328-352-1277	Sequence 1277, Ap
C 19	32.8	2.4	414	4 US-09-806-708B-22	Sequence 22, Appl
C 20	32.8	2.4	1416	4 US-09-902-540-872	Sequence 872, App
C 21	32.6	2.4	10813	4 US-08-956-171B-364	Sequence 364, App
C 22	32.6	2.4	10813	4 US-08-781-986A-364	Sequence 364, App
C 23	32.4	2.4	630	4 US-09-248-786A-5221	Sequence 5221, Ap
C 24	32.4	2.4	28960	4 US-09-949-016-11978	Sequence 11978, A
C 25	32.4	2.4	30054	4 US-09-949-016-16100	Sequence 16100, A
C 26	32.2	2.3	978	3 US-09-267-031-13	Sequence 13, Appl
C 27	32.2	2.3	89843	4 US-09-949-016-12346	Sequence 12346, A

28	32.2	2.3	89844	4 US-09-949-016-13656	Sequence 13656, A
29	32	2.3	988	3 US-08-793-634B-7	Sequence 7, Appl1
30	32	2.3	1281	3 US-09-359-070-1	Sequence 1, Appl1
31	32	2.3	1300	2 US-08-440-845D-8	Sequence 8, Appl1
32	32	2.3	1300	2 US-08-868-458-8	Sequence 8, Appl1
33	32	2.3	1300	3 US-09-134-262-1	Sequence 1, Appl1
C 34	32	2.3	10445	4 US-09-949-016-12311	Sequence 12311, A
C 35	32	2.3	10445	4 US-09-949-016-12927	Sequence 12927, A
C 36	32	2.3	42741	4 US-09-949-016-11857	Sequence 11857, A
C 37	31.8	2.3	69874	4 US-09-949-016-12361	Sequence 12361, A
38	31.8	2.3	69874	4 US-09-949-016-13049	Sequence 13049, A
39	31.8	2.3	94077	4 US-09-949-016-13635	Sequence 13635, A
C 40	31.8	2.3	278866	4 US-09-949-016-13922	Sequence 13922, A
C 41	31.8	2.3	278866	4 US-09-949-016-13923	Sequence 13923, A
C 42	31.8	2.3	278866	4 US-09-949-016-13924	Sequence 13924, A
C 43	31.8	2.3	278866	4 US-09-949-016-13925	Sequence 13925, A
C 44	31.8	2.3	278866	4 US-09-949-016-13926	Sequence 13926, A
C 45	31.8	2.3	278866	4 US-09-949-016-14699	Sequence 14699, A

ALIGNMENTS

RESULT 1
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
CLASSIFICATION: 435
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpc-F18
US-08-232-463-14
Query Match 3.1%, Score 42.2, DB 1, Length 7218;


```

RESULT 2
US-09-007-005-17
; Sequence 17, Application US/0907005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihue
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

```

```

Db      1  RGRGRGARCAARARURURARCBURARURURURARCBARARURURARCBARARURGRNRN 60
Qy      1095  TGTGCTTTCGCGCTGCTACCTTACACATCATCGGGGCTCAACACGATACACTTACGCTTCA 1154
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      61  RSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRN 120
Qy      1155  AGAGTGGCGCATGGAAGTACTACCCGATCGAACCCGGAAAGCGTTCAACATCGGCGT 1214
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      121  RSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRN 180
Qy      1215  ACCTGATACATAGCTACATAGAAACATTAAGAGCTTGAAGAAC 1260
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      181  RSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRN 226

RESULT 3
US-09-244-796-17
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihc
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350007
; CURRENT APPLICATION NUMBER: US/09/244,796
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 09/007,005
; EARLIER FILING DATE: 1998-01-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; NAME/KEY: misc_feature
; LOCATION: (1)..(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

Query Match      2.7%; Score 37.6; DB 3; Length 289;
Best Local Similarity 4.0%; Pred. No. 0.035;
Matches 9; Conservative 103; Mismatches 114; Indels 0; Gaps 0;

Qy      1035  GGGGCAGAGCAACTATTTTCAATGAATCAATCAATTCGGTAGAGTTCCAA 1094
           :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      1  RGRGRGARCAARARURURARCBURARURURURARCBARARURURARCBARARURGRNRN 60
Qy      1095  TGTGCTTTCGCGCTGCTACCTTACACATCATCGGGGCTCAACACGATACACTTACGCTTCA 1154
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      61  RSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRN 120
Qy      1155  AGAGTGGCGCATGGAAGTACTACCCGATCGAACCCGGAAAGCGTTCAACATCGGCGT 1214
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      121  RSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRN 180
Qy      1215  ACCTGATACATAGCTACATAGAAACATTAAGAGCTTGAAGAAC 1260
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      181  RSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRN 226

RESULT 4
US-09-621-976-8976/C
; Sequence 8976, Application US/09621976
; Patent No. 6639063

```

```

; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8976
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-621-976-8976

Query Match          2.6%; Score 36; DB 4; Length 399;
Best Local Similarity 11.4%; Pred. No. 0.16;
Matches 41; Conservative 164; Mismatches 154; Indels 1; Gaps 1;

Qy 669 AGGATGGGGGCTAAGTCGGCATTTGAAGGATTAAGGGGGTGTGAAGTGTGTCA 728
Db 379 ACGTATGAGCATCAATCAGCCAAAGKMKWMSGAMKMRGTWYWMYTSKSYRRYKTS 320
Qy 729 AAAGAGGTGCGATTGCTATACACGCGCTAAGCAGGTGGGCTAGCAG-CTGTCTGACGC 787
Db 319 AAMGRATKKRKTATKKTWMMGGGKGGSTYMAKRSRSGTGRMSYRAAMRSGSKWGG 260
Qy 788 TGTGAATAACTGCTAGTGTAGTATGTCCAACTAATGTCAGAGATGCAATGCTGAT 847
Db 259 SYVMAGYRSRMSRWSYAMRKKMKTKMSRSGSTGYVAMMYKMSCTSRKMMY 200
Qy 848 TGGGTAAATGGGCATAGTGTAGTGGCCGAAAACGTTTGTAGTCTAGTTAAAGGA 907
Db 199 KKRKKMRKCTSTKRTCTGTYKCMFAYYTKRKRKTATWYTYKYSYMSMKTKTMRKT 140
Qy 908 AGCTGAAGCTGAACCTGTGCAAAATAGCCTGTGTAATCAACGTTGATTAACCAATT 967
Db 139 AYYTKRKRKTTRTKTWTCTMCMKCTTYMAGTMMYRYYAKRASKRCTWSTTCY 80
Qy 968 CAGTCGTCAAGGGTCTCTGATATGCTGAGACCTTCCTGTGCGATTGTGGGTAATCT 1027
Db 79 CMKYMAKKCWSYMSMSMKWMSMKWYKTYYYYYYMMKMSKMTYMSWMSCYARKCWRTY 20

RESULT 5
US-09-621-976-15639/c
; Sequence 15639, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15639
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-621-976-15639

Query Match          2.6%; Score 35.2; DB 4; Length 505;
Best Local Similarity 15.6%; Pred. No. 0.35;
Matches 28; Conservative 82; Mismatches 70; Indels 0; Gaps 0;

Qy 616 ATGATGCCAATCTTAAAGAGGAGTGTGAGACATTTTAAAGGATCAATATAGCAATT 675
Db 344 MTGRKGSMTKRKMMYTGMMWTSYKCTKTKGKYTGMSKSKTRMTCTSMRKYMMMSGCM 285
```

```

Qy 676 GGGCGCTAAGTCGCGCATTTGAAGGATTAAGGGGGGTGTGAAGTGTGTCAAAAGAG 735
Db 284 ASMKSWARSWYSWMAICWMSASAYRABRSMYCARSRMWRAGAGWWRARBRGKRARGKS 225
Qy 736 GTCGATTGGCTATATCCAGCGCGCTAAGCAGGTGGGCTAGCAGCTGTCTGCAGCTGATA 795
Db 224 SMNSKMSWRMSXGAKRMCBRWMSCRMSYSICMSGKSCMSCRGTCAKMRBYARAYAKRYA 165

RESULT 6
US-09-949-016-16369/c
; Sequence 16369, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CD001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16369
; LENGTH: 136917
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-16369

Query Match          2.5%; Score 35; DB 4; Length 136917;
Best Local Similarity 55.3%; Pred. No. 19;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 1210 GGCCTAGCTAGATACATAGAGTCAATAGAAAGAGAGCTTGAAGAACATTCAAAT 1269
Db 106817 GTCCTAGCTTCTTCTTAAACAAATACCAATATCCAGGGCTAGTAGACCTTGAAT 106758
Qy 1270 CCTAAGGCTCTCTCTCTTCTTGTGCATCAGATCAAGAAATCATACACTAAACAGGACTC 1329
Db 106757 ATACAGAAACATCTTCTCTTAAGAGAAAACCTCAGAGATTAATTAATCAATAGAAACC 106698

Qy 1330 TTT 1332
Db 106697 ATT 106695

RESULT 7
US-09-621-976-2813
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-621-976-2813

Query Match          2.5%; Score 35; DB 4; Length 832;
Best Local Similarity 55.3%; Pred. No. 19;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 1210 GGCCTAGCTAGATACATAGAGTCAATAGAAAGAGAGCTTGAAGAACATTCAAAT 1269
Db 106817 GTCCTAGCTTCTTCTTAAACAAATACCAATATCCAGGGCTAGTAGACCTTGAAT 106758
Qy 1270 CCTAAGGCTCTCTCTTCTTGTGCATCAGATCAAGAAATCATACACTAAACAGGACTC 1329
Db 106757 ATACAGAAACATCTTCTCTTAAGAGAAAACCTCAGAGATTAATTAATCAATAGAAACC 106698

Qy 1330 TTT 1332
Db 106697 ATT 106695
```


Sequence 205, Application US/09618166
Patent No. 6583112
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
Yu, Chang-Hu
Oshima, Junko
Mulligan, John T.
Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSER: Seed Intellectual Property Law Group
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/618,166
APPLICATION NUMBER: US/09/618,166
FILING DATE: 17-Jul-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McMaisters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 240052.419C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 205:
SEQUENCE CHARACTERISTICS:
LENGTH: 4792 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 145..4347
SEQUENCE DESCRIPTION: SEQ ID NO: 205:
US-09-618-166-205
Query Match 2.4%; Score 33.4; DB 4; Length 4792;
Best Local Similarity 49.7%; Pred. No. 6.8; Mismatches 86; Indels 0; Gaps 0;
Matches 85; Conservative 0;
213 ATGACACCCACCCGCAAGCAATCCTGGAGTTCAAGAGGGTGTCTTACGGCATTTA 272
1164 ATGGACCCGCAAGCAATCCTGGAGTTCAAGAGGGTGTCTTACGGCATTTA 1223
QY 273 GGTATAGATGCGATAGCGTTTGAAGCTGAAGCTGATTCGACAGCATGAGACAAG 332
DB 1224 ACTGAAGCAAGAAAAGTGAATCTGAATAATGAAGAAATGAAATCTGTTGAGAGAAA 1283
QY 333 AAAATACACGGTTGATGCGTTCCCGTCTTAACTAAATGATATCCACA 383
DB 1284 TATGAAAAGAACTGTGTATCTTCTAGTATTTTCAGAAAATGAATCTCCACA 1334

RESULT 14
US-09-127-670-5
Sequence 5, Application US/09127670
Patent No. 6228583
GENERAL INFORMATION:
APPLICANT: Massachusetts Institute for Technology
APPLICANT: Leonard P. Guarente
APPLICANT: David A. Sinclair
APPLICANT: David B. Lombard
TITLE OF INVENTION: ASSAYS FOR COMPOUNDS WHICH EXTEND LIFE

TITLE OF INVENTION: SPAN
FILE REFERENCE: MIT-7720PA
CURRENT APPLICATION NUMBER: US/09/127,670
CURRENT FILING DATE: 1998-07-31
EARLIER APPLICATION NUMBER: US 60/054,629
EARLIER FILING DATE: 1997-08-04
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 6476
TYPE: DNA
ORGANISM: Murine
FEATURE:
NAME/KEY: CDS
LOCATION: (229)...(4432)
US-09-127-670-5
Query Match 2.4%; Score 33.4; DB 3; Length 6476;
Best Local Similarity 49.7%; Pred. No. 8.4; Mismatches 86; Indels 0; Gaps 0;
Matches 85; Conservative 0;
213 ATGACACCCACCCGCAAGCAATCCTGGAGTTCAAGAGGGTGTCTTACGGCATTTA 272
DB 1248 ATGGACCCGCAAGCAATCCTGGAGTTCAAGAGGGTGTCTTACGGCATTTA 1307
QY 273 GGTATAGATGCGATAGCGTTTGAAGCTGAAGCTGATTCGACAGCATGAGACAAG 332
DB 1308 AGTGAAGCAAGAAAAGTGAATCTGAATAATGAAGAAATGAAATCTGTTGAGAGAAA 1367
QY 333 AAAATACACGGTTGATGCGTTCCCGTCTTAAAGTATATCCACA 383
DB 1368 TATGAAAAGAACTGTGTATCTTCTAGTATTTTCAGAAAATGAATCTCCACA 1418

RESULT 15
US-08-781-891-208
Sequence 208, Application US/08781891
Patent No. 6090820
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-Hu
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSER: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/781,891
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620cendburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 208:
SEQUENCE CHARACTERISTICS:
LENGTH: 16442 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-781-891-208

Query Match 2.4%; Score 33.4; DB 3; Length 16442;
Best Local Similarity 49.7%; Pred. No. 16;
Matches 85; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

OY 213 ATGACACCCACCCCGCAACGGAATCCTGAGTTCAAGAGGGTGCCTACGGCATTTA 272
13744 ATGGACCCAGAACTTGACAGTTTAGTGAAGCAAGAGGGTTGATGTAATTTAGAAATCA 13803
OY 273 GGTATAGATGCGATAGGGTTTGACGTAAGCTGAAAGCTGATTACGACATGAGACAAAG 332
13804 AGTGAAGCAAGAAAAGGTGAATCTGAAAATGAAATAGAAAGATAATCTGTTGAGAGAAAG 13863
OY 333 AAAATACAAAGGTGTATGCGTTCCCGTCTTACTAAAGTGATATCCAAAG 383
13864 TATGGAAGAACTGTGTGATTCCTAGTATTCAGAAAATGAATCCAAAG 13914
DB

Search completed: February 9, 2005, 13:44:38
Job time : 275 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 12:28:25 ; Search time 763 Seconds
(without alignments)
10395.036 Million cell updates/sec

Title: US-10-070-386-1
Perfect score: 1378
Sequence: 1 gtgcagcgggtggtgatcat.....accatcatcaactaaca 1378

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4313806 seqs, 2877871033 residues

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36.4	2.6	600	US-10-123-155-462	Sequence 462, App
2	36.4	2.6	600	US-10-146-731-462	Sequence 462, App
3	36.4	2.6	600	US-10-140-472-462	Sequence 462, App
4	36.4	2.6	600	US-10-141-761-462	Sequence 462, App
5	36.4	2.6	600	US-10-142-885-462	Sequence 462, App
6	36.4	2.6	600	US-10-158-790-462	Sequence 462, App
7	36.4	2.6	600	US-10-137-871-462	Sequence 462, App
8	36.4	2.6	600	US-10-140-923-462	Sequence 462, App
9	36.4	2.6	600	US-10-141-756-462	Sequence 462, App
10	36.4	2.6	600	US-10-141-759-462	Sequence 462, App
11	36.4	2.6	600	US-10-140-805-462	Sequence 462, App

12	36.4	2.6	600	US-10-140-864-462	Sequence 462, App
13	36.4	2.6	600	US-10-142-426-462	Sequence 462, App
14	36.2	2.6	2672	US-09-775-938A-34	Sequence 34, App1
15	36	2.6	3673778	US-10-312-841-1	Sequence 1, App1
16	35.8	2.6	134738	US-10-087-192-502	Sequence 502, App1
17	35.4	2.6	993	US-09-898-837A-8	Sequence 8, App1
18	35	2.5	458	US-10-184-644-498	Sequence 498, App
19	35	2.5	458	US-10-184-634-498	Sequence 498, App
20	35	2.5	537	US-10-027-632-774500	Sequence 274500,
21	35	2.5	537	US-10-027-632-774500	Sequence 274500,
22	35	2.5	6971	US-10-311-455-1210	Sequence 1210, App
23	35	2.5	207433	US-10-277-216-5	Sequence 5, App11
24	35	2.5	207433	US-10-126-022-5	Sequence 36862, A
25	34.8	2.5	4416	US-10-369-493-16862	Sequence 98, App1
26	34.4	2.5	8805	US-10-257-166-98	Sequence 154, App
27	34.4	2.5	32874	US-10-052-482-154	Sequence 1796, App
28	34.2	2.5	312	US-10-674-124A-7196	Sequence 160, App
29	34.2	2.5	90435	US-10-332-696-160	Sequence 39026, A
30	34	2.5	1591	US-10-437-963-39026	Sequence 807, App1
31	34	2.5	6283	US-10-311-455-807	Sequence 73, App1
32	34	2.5	6283	US-10-257-166-73	Sequence 73, App1
33	34	2.5	339234	US-10-322-696-73	Sequence 16934,
34	33.8	2.5	818	US-10-027-632-166934	Sequence 16934,
35	33.8	2.5	818	US-10-027-632-166934	Sequence 16934,
36	33.8	2.5	818	US-10-027-632-166934	Sequence 16934,
37	33.8	2.5	818	US-10-027-632-166934	Sequence 16934,
38	33.6	2.4	25179	US-10-322-281-277	Sequence 277, App
39	33.4	2.4	403	US-09-864-761-5003	Sequence 5003, App
40	33.4	2.4	737	US-10-425-115-65924	Sequence 65924, A
41	33.4	2.4	1024	US-10-123-155-198	Sequence 198, App
42	33.4	2.4	1024	US-10-146-731-198	Sequence 198, App
43	33.4	2.4	1024	US-10-140-472-198	Sequence 198, App
44	33.4	2.4	1024	US-10-141-761-198	Sequence 198, App
45	33.4	2.4	1024	US-10-142-885-198	Sequence 198, App

ALIGNMENTS

RESULT 1
US-10-123-155-462
; Sequence 462, Application US/10123155
; Publication No. US20030068794A1
GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Laureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C30
CURRENT APPLICATION NUMBER: US/10/123,155
PRIORITY FILING DATE: 2002-04-15
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 462
LENGTH: 600
TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-155-462

```
Query Match      2.6%; Score 36.4; DB 14; Length 600;
Best Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;

QY 608 AGACCTAATGATGCGACATCTTAAGACGAGATCTGAGACATTTCTTAAGCATCCATA 667
DB 1 MRSCLMRCRHLISQGVMSLLAVLVPFLPALPSFIKEPQTPKSRHQRTEINKERSLQSLA 60
QY 668 TAGGCATTGGCGCGTAACTCGGCGATTGAAGAGATGAAGGGGGTGAAGTGGTGTCT 727
DB 61 KPKQAPFRARRTTIYABAPENNALNTQPKATHTGDRKEANQAPPEBDKVPHTAQ 120
QY 728 AAAAGAGGTGATGGCTATACACCGCTAAGCAGGTGGCTAGCAGCTGTCTGCAGC 787
DB 121 RAAMKSPKEKTMVNTLSPRQDAGMASGRTEAQSWSQDTKTTQNGGQTRKLTASRTV 180
QY 788 TGTGAATPACGTCATCTTGTAGTATGCCACCTAATGTGAGCAGATGCAAAATGCTAT 847
DB 181 SEKHQKAAATTAATKTLIPKSHRMLAPTGAVSTRTRQKGVTTAVIPPKKKPQATPPAPF 240
QY 848 TGGGTTAAATGGCATGTAGTAGTGCCGAAACAGCTTATGATCTAGTTAAAGGA 907
DB 241 QSPTRQNRQLKAANFKSEPRWDPEEKYSFEIGLQTTCPDVSXIKAKSLMLQKLPJN 300
QY 908 AGCTGAAAAGCTGAACCTGTGAGAAATPAACTGTTGAATACACCTTGATPAAACCAATT 967
DB 301 LTLFLDSHFHQSEWDRLEHFAFPFGFMELNYSLVQKVTTRPPVPOQULLASLPAGSL 360
QY 968 CAGTGTCAAGGGTGTCTCGTATATCTGAGCTTCCCTGCGATTTGGGTAATCTATT 1027
DB 361 RCITCAVAVNGGILNNSHMGDEIDSHDYVFLSGALIKGYEDVGRTRSFYGTAFSLTQ 420
QY 1028 TCATAGTGGGCGAGATCACTCTATTTCATTTGAATCTAACTATTTCTGGTAGAG 1087
DB 421 SLLILGNKGFKNVPLGKDVRYLHFLBGRIDYEWLEBALMNTVMSKNLFWFRRHQEAER 480
QY 1088 TTCTCAATGCTCTT 1101
DB 481 EALHMDRYLLHPD 494

RESULT 2
US-10-146-731-462
; Sequence 462, Application US/10146731
; Publication No. US20030129692A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C323
; CURRENT APPLICATION NUMBER: US/10/146,731
; FILE REFERENCE: P3330R1C168
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 462
; LENGTH: 600
```

```
TYPE: PRT
; ORGANISM: Homo Sapien
US-10-146-731-462

Query Match      2.6%; Score 36.4; DB 15; Length 600;
Best Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;

QY 608 AGACCTAATGATGCGACATCTTAAGACGAGATCTGAGACATTTCTTAAGCATCCATA 667
DB 1 MRSCLMRCRHLISQGVMSLLAVLVPFLPALPSFIKEPQTPKSRHQRTEINKERSLQSLA 60
QY 668 TAGGCATTGGCGCGTAACTCGGCGATTGAAGAGATGAAGGGGGTGAAGTGGTGTCT 727
DB 61 KPKQAPFRARRTTIYABAPENNALNTQPKATHTGDRKEANQAPPEBDKVPHTAQ 120
QY 728 AAAAGAGGTGATGGCTATACACCGCTAAGCAGGTGGCTAGCAGCTGTCTGCAGC 787
DB 121 RAAMKSPKEKTMVNTLSPRQDAGMASGRTEAQSWSQDTKTTQNGGQTRKLTASRTV 180
QY 788 TGTGAATPACGTCATCTTGTAGTATGCCACCTAATGTGAGCAGATGCAAAATGCTGAT 847
DB 181 SEKHQKAAATTAATKTLIPKSHRMLAPTGAVSTRTRQKGVTTAVIPPKKKPQATPPAPF 240
QY 848 TGGGTTAAATGGCATGTAGTAGTGCCGAAACAGCTTATGATCTAGTTAAAGGA 907
DB 241 QSPTRQNRQLKAANFKSEPRWDPEEKYSFEIGLQTTCPDVSXIKAKSLMLQKLPJN 300
QY 908 AGCTGAAAAGCTGAACCTGTGAGAAATPAACTGTTGAATACACCTTGATPAAACCAATT 967
DB 301 LTLFLDSHFHQSEWDRLEHFAFPFGFMELNYSLVQKVTTRPPVPOQULLASLPAGSL 360
QY 968 CAGTGTCAAGGGTGTCTCGTATATCTGAGCTTCCCTGCGATTTGGGTAATCTATT 1027
DB 361 RCITCAVAVNGGILNNSHMGDEIDSHDYVFLSGALIKGYEDVGRTRSFYGTAFSLTQ 420
QY 1028 TCATAGTGGGCGAGATCACTCTATTTCATTTGAATCTAACTATTTCTGGTAGAG 1087
DB 421 SLLILGNKGFKNVPLGKDVRYLHFLBGRIDYEWLEBALMNTVMSKNLFWFRRHQEAER 480
QY 1088 TTCTCAATGCTCTT 1101
DB 481 EALHMDRYLLHPD 494

RESULT 3
US-10-140-472-462
; Sequence 462, Application US/10140472
; Publication No. US2003013888A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C168
; CURRENT APPLICATION NUMBER: US/10/140,472
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See File Wrapper or Palm
```

```
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO: 462
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-472-462
```

```
Query Match      2.6%; Score 36.4; DB 15; Length 600;
Beet Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;
```

```
OY 608 AGACCTACATGATGCCACATCTTAAAGCAGGATCTGACACATTTCTAAGGCATCCATA 667
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1 MRSCLMRCHLSQGVOMSLLAIVLFPLPSPFIKEPQTPSRHORENTERKESLQSLA 60
OY 668 TAGGATTGGGCGCTAAGTCGGCATTTGAAGAGATTAAGGGGGGTGGAAGGTGTGTC 727
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 61 KPKSQAPTRARRTTYAEPAPENNALNTQTPKHTTGDGRKANQAPPEODKVPHTAQ 120
OY 728 AAAAGAGGTCGATTGGCTATACAGCCGCTAAGCAGGTGGCTAGCAGCTGTGCAGC 787
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 121 RAAMWSPEKEMVNTLSPRGDAGMASGRTESQMSQDPYKTGGNGGQTRKLTASRTV 180
OY 788 TGTGAATAAGCTCACTTGCTTAGGTATGTCACCTAATGTGCAGAGATGCAAAATGCTGAT 847
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 181 SEKHGKATTAATKTLIPKSHMLAPTGAVSTRTRQKGVTTAVIPPEKKPQATPPAPF 240
OY 848 TGGGTTAAATGGGATGTAGTGTAGTCCGCAAAAACAGCTTAATCTAGTTAAAGGA 907
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 241 QSPITQRNQRLLKANFKSEPRWDPEEKYSFEIGLQTTCPDSVKIKAKSLMLQKLPFN 300
OY 908 AGCTGAAGCTGAACCTGTCAAGAAATAAGCTGTGAATAACAAGTTGATTAACCAATT 967
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 301 LTLFLDSHRHFNQSEMDRLHFAFPFGFMELNYSLVQKVTFRPPVPOOULLASIPAGSL 360
OY 968 CAGTCGTAAGGTCGTCTGATATGCTGGAGCTCCCTGCGCATTTGGGGTAACTATT 1027
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 361 RCITCAVANGGILNNSHMGDEIDSHDYVFRLSGALIGYEDVDGTRISFYGTFASLTQ 420
OY 1028 TCATAGTGGGCGAGAAATGCAACTCTATTTCATTAATGAATCTAATCTATTTGGGTAGAG 1087
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 421 SLLILNRRGFKNVPLGKQVRYLHPLFGTRDYEMLEALLMNTQVMSKNLFWFRHRQBAFR 480
OY 1088 TTCTCAATGTCCTT 1101
Db 481 EALHMDRYLLHPD 494
```

```
RESULT 4
US-10-141-761-462
; Sequence 462, Application US/10141761
; Publication No. US20030148432A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330RIC198
```

```
; CURRENT APPLICATION NUMBER: US/10/141,761
; PRIOR FILING DATE: 2002-05-08
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO: 462
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-761-462
```

```
Query Match      2.6%; Score 36.4; DB 15; Length 600;
Beet Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;
```

```
OY 608 AGACCTACATGATGCCACATCTTAAAGCAGGATCTGACACATTTCTAAGGCATCCATA 667
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1 MRSCLMRCHLSQGVOMSLLAIVLFPLPSPFIKEPQTPSRHORENTERKESLQSLA 60
OY 668 TAGGATTGGGCGCTAAGTCGGCATTTGAAGAGATTAAGGGGGGTGGAAGGTGTGTC 727
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 61 KPKSQAPTRARRTTYAEPAPENNALNTQTPKHTTGDGRKANQAPPEODKVPHTAQ 120
OY 728 AAAAGAGGTCGATTGGCTATACAGCCGCTAAGCAGGTGGCTAGCAGCTGTGCAGC 787
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 121 RAAMWSPEKEMVNTLSPRGDAGMASGRTESQMSQDPYKTGGNGGQTRKLTASRTV 180
OY 181 SEKHGKATTAATKTLIPKSHMLAPTGAVSTRTRQKGVTTAVIPPEKKPQATPPAPF 240
OY 241 QSPITQRNQRLLKANFKSEPRWDPEEKYSFEIGLQTTCPDSVKIKAKSLMLQKLPFN 300
OY 301 LTLFLDSHRHFNQSEMDRLHFAFPFGFMELNYSLVQKVTFRPPVPOOULLASIPAGSL 360
OY 361 RCITCAVANGGILNNSHMGDEIDSHDYVFRLSGALIGYEDVDGTRISFYGTFASLTQ 420
OY 420 TCATAGTGGGCGAGAAATGCAACTCTATTTCATTAATGAATCTAATCTATTTGGGTAGAG 1087
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1028 TCATAGTGGGCGAGAAATGCAACTCTATTTCATTAATGAATCTAATCTATTTGGGTAGAG 1087
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 421 SLLILNRRGFKNVPLGKQVRYLHPLFGTRDYEMLEALLMNTQVMSKNLFWFRHRQBAFR 480
OY 480 TTCTCAATGTCCTT 1101
Db 1101 EALHMDRYLLHPD 494
```

```
RESULT 5
US-10-142-885-462
; Sequence 462, Application US/10142885
; Publication No. US20030157604A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
```

```

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P33081C248
; CURRENT APPLICATION NUMBER: US/10/142,885
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 462
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-885-462

```

```

Query Match      2.6%; Score 36.4; DB 16; Length 600;
Best Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;

```

```

QY 608 AGACCTATGATGCGCATCTTAAGAGGAGATCTGAGACATTTCTAAGGCATCCATA 667
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 MRSCIMRCRHLSQGVOMSLLAIVLFPLPSPFIKEPQTPSRHQRTEINIKERSLSGLA 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 668 TAGGCATTGGCGCGCTAAGTCCGATTTGAAGAGATTAAGGGGGGTGTAAGTGTGTC 727
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 61 KPKSQAPTRARRTTYIAPAPENNALNTQTPKATHTTDRKEANQAPPEEQDKVPHTAQ 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 728 AAAAGAGAGTGGATGGCTATACAGCCGCTAAGCAGGTGGCTAGCAGCTGTGCAGC 787
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 121 RAAMKSPKEKTMVNTLSPRQDAGMASGRTEAQSWKSQDFTKTTQNGGQTRKLTASRTV 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 788 TGTGAATTAACGTCACTTGATAGTATGTCACCTTAATGTGAGCAGATGCAAAATGCTAT 847
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 181 SEKHQKATATTAATKLIPRSQHRMLAPTGAVSTRTRQKQVTTAVIPPKKKPOATPPAPF 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 848 TGGGTTAAATGGCGATGATAGTAGTGCCGAAAACAGTTTAACTAGTTAAAGGA 907
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 241 QSPTRQNRQRLKAANFKSEPRWDPEEKYSFEIGLQTTCPDSVKIKAKSKMLQKLFPLN 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 908 AGCTGAAGCTGAACCTGTGAGAAATTAAGCCTGTGGAATACACGTTGATTAACCAATT 967
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 301 LTLFLDSHNFQSEMDRLNHFAPRFGFMEINYSLVQKVTTRFPVPOQQLLASLPAGSL 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 968 CAGTGTCAAGGGTCTCTGATATGCTGAGCTTCCCTGTGCAATGTGGGTAATACTATT 1027
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 361 RCITCAVAVNGGILNNSHMGDEIDSHDYFRLSGALIKGYEQDVGTSTRSPYGFATFSLTQ 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1028 TCATAGTGGGCGAGAACTGATCTATTTCATTTGAATCTAACTATTCGGTAGAG 1087
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 421 SLILIGNRGFKNVPLGKQVRYLHLEGTTRDYEMLEALLMNTVMSKNLFWFHRHPOEAFR 480
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1088 TTCTCAATGCTCTT 1101
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 481 EALHMDRYLLHHPD 494
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 6
US-10-158-790-462
; Sequence 462, Application US/10158790
; Publication No. US20030180879A1

```

```

; GENERAL INFORMATION:

```

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel

```

```

; APPLICANT: Matanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P33081C448
; CURRENT APPLICATION NUMBER: US/10/158,790
; CURRENT FILING DATE: 2002-05-30
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 462
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-158-790-462

```

```

Query Match      2.6%; Score 36.4; DB 16; Length 600;
Best Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;

```

```

QY 608 AGACCTATGATGCGCATCTTAAGAGGAGATCTGAGACATTTCTAAGGCATCCATA 667
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 MRSCIMRCRHLSQGVOMSLLAIVLFPLPSPFIKEPQTPSRHQRTEINIKERSLSGLA 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 668 TAGGCATTGGCGCGCTAAGTCCGATTTGAAGAGATTAAGGGGGGTGTAAGTGTGTC 727
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 61 KPKSQAPTRARRTTYIAPAPENNALNTQTPKATHTTDRKEANQAPPEEQDKVPHTAQ 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 728 AAAAGAGAGTGGATGGCTATACAGCCGCTAAGCAGGTGGCTAGCAGCTGTGCAGC 787
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 121 RAAMKSPKEKTMVNTLSPRQDAGMASGRTEAQSWKSQDFTKTTQNGGQTRKLTASRTV 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 788 TGTGAATTAACGTCACTTGATAGTATGTCACCTTAATGTGAGCAGATGCAAAATGCTAT 847
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 181 SEKHQKATATTAATKLIPRSQHRMLAPTGAVSTRTRQKQVTTAVIPPKKKPOATPPAPF 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 848 TGGGTTAAATGGCGATGATAGTAGTGCCGAAAACAGTTTAACTAGTTAAAGGA 907
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 241 QSPTRQNRQRLKAANFKSEPRWDPEEKYSFEIGLQTTCPDSVKIKAKSKMLQKLFPLN 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 908 AGCTGAAGCTGAACCTGTGAGAAATTAAGCCTGTGGAATACACGTTGATTAACCAATT 967
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 301 LTLFLDSHNFQSEMDRLNHFAPRFGFMEINYSLVQKVTTRFPVPOQQLLASLPAGSL 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 968 CAGTGTCAAGGGTCTCTGATATGCTGAGCTTCCCTGTGCAATGTGGGTAATACTATT 1027
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 361 RCITCAVAVNGGILNNSHMGDEIDSHDYFRLSGALIKGYEQDVGTSTRSPYGFATFSLTQ 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1028 TCATAGTGGGCGAGAACTGATCTATTTCATTTGAATCTAACTATTCGGTAGAG 1087
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 421 SLILIGNRGFKNVPLGKQVRYLHLEGTTRDYEMLEALLMNTVMSKNLFWFHRHPOEAFR 480
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1088 TTCTCAATGCTCTT 1101
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 481 EALHMDRYLLHHPD 494
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 7
US-10-137-871-462
; Sequence 462, Application US/10137871
; Publication No. US20030207350A1

```

```

; GENERAL INFORMATION:

```

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven

```

```

; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C153
; CURRENT APPLICATION NUMBER: US/10/137,871
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 462
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-137-871-462

```

```

Query Match      2.6%; Score 36.4; DB 17; Length 600;
Best Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;

```

```

QY 608 AGACCTACATGATGCCATCTTAAGCAGGAGTGTGACATTTCTTAAGGCATCCATA 667
DB 1 MRSCLMRCHLSQGVQWMLLAVLFPLPSPFIKEPQTKPSRHQRTENIKERSLQSLA 60
QY 668 TAGGCATTGGCGCGCTTAAGCGCATTTGAAGAGATTAAGGGGGGTGAAAGTGTGTCTC 727
DB 61 KRKSGAPTRARRTTIYAPAPENNALNTQTPKHTTGDREKANQAPREBDKVPHTAQ 120
QY 728 AAAAGAGGTGATGGCTATACACAGCGCTAAGCAGGTGGGTGACAGCTGTGACGC 787
DB 121 RAAMWSPEKEXTMVTLSPRGODAGMASGRTEAOSKSDPTTQNGGQTRKLTASRTV 180
QY 788 TGTGAATAAGCTCACTTGCTAGTATGTCACCTAATGTCAGAGATGCAATGCTGAT 847
DB 181 SEKHGKATTAATKTLIPKQNHMLAPTVGAVSTRTRQKGVTTAVIPPEKKEPQATPPAPF 240
QY 848 TGGGTTAAATGGCGATGATGATGATGTCGCCAAAACAGCTTAATGCTAATGTTAAAGGA 907
DB 241 QSPTTQRNORLKAANFKSEPRWDPEBKYSFEIGLQTTCPDSVKIKASKSLMLQKLPFN 300
QY 908 AGCTGAAGCTGAACCTGTGAGAATAAAGCTGTGGAATACAGTTGATACCAATT 967
DB 301 LTLFLDSRHFNOSEWDRLEHFAAPPFGEFMEIANTSLVQKVTRPPVPQOQLLASLPAGSL 360
QY 968 CAGTGTCAAGGCTGCTGATATGCTGAGCTTCCCTGCGCATTTGGGGTAACTATT 1027
DB 361 RCITCAVANGGILNNSHMGGEIDSHDVFRISGALIKGYEDVGTRTSFYGTAFSLTQ 420
QY 1028 TCATAGTGGGCGAGAACTCAATCTTATTTCAATGAATCAACTTAATCTTGGGTAGAG 1087
DB 421 SLLILGNRGFKNVPKGVRYLHFLGTRDYEMLEALLMQTVMSKNLFWFHRHROEAFR 480
QY 1088 TTCTCAATGCTCTT 1101
DB 481 EALHMDRYLLHPD 494

```

```

RESULT 8
US-10-140-923-462

```

```

; Sequence 462, Application US/10140923
; Publication No. US20030207355A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey

```

```

; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C188
; CURRENT APPLICATION NUMBER: US/10/140,923
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 462
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-923-462

```

```

Query Match      2.6%; Score 36.4; DB 17; Length 600;
Best Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;

```

```

QY 608 AGACCTACATGATGCCATCTTAAGCAGGAGTGTGACATTTCTTAAGGCATCCATA 667
DB 1 MRSCLMRCHLSQGVQWMLLAVLFPLPSPFIKEPQTKPSRHQRTENIKERSLQSLA 60
QY 668 TAGGCATTGGCGCGCTTAAGCGCATTTGAAGAGATTAAGGGGGGTGAAAGTGTGTCTC 727
DB 61 KRKSGAPTRARRTTIYAPAPENNALNTQTPKHTTGDREKANQAPREBDKVPHTAQ 120
QY 728 AAAAGAGGTGATGGCTATACACAGCGCTAAGCAGGTGGGTGACAGCTGTGACGC 787
DB 121 RAAMWSPEKEXTMVTLSPRGODAGMASGRTEAOSKSDPTTQNGGQTRKLTASRTV 180
QY 788 TGTGAATAAGCTCACTTGCTAGTATGTCACCTAATGTCAGAGATGCAATGCTGAT 847
DB 181 SEKHGKATTAATKTLIPKQNHMLAPTVGAVSTRTRQKGVTTAVIPPEKKEPQATPPAPF 240
QY 848 TGGGTTAAATGGCGATGATGATGATGTCGCCAAAACAGTTAATGCTAATGTTAAAGGA 907
DB 241 QSPTTQRNORLKAANFKSEPRWDPEBKYSFEIGLQTTCPDSVKIKASKSLMLQKLPFN 300
QY 908 AGCTGAAGCTGAACCTGTGAGAATAAAGCTGTGGAATACAGTTGATACCAATT 967
DB 301 LTLFLDSRHFNOSEWDRLEHFAAPPFGEFMEIANTSLVQKVTRPPVPQOQLLASLPAGSL 360
QY 968 CAGTGTCAAGGCTGCTGATATGCTGAGCTTCCCTGCGCATTTGGGGTAACTATT 1027
DB 361 RCITCAVANGGILNNSHMGGEIDSHDVFRISGALIKGYEDVGTRTSFYGTAFSLTQ 420
QY 1028 TCATAGTGGGCGAGAACTCAATCTTATTTCAATGAATCAACTTAATCTTGGGTAGAG 1087
DB 421 SLLILGNRGFKNVPKGVRYLHFLGTRDYEMLEALLMQTVMSKNLFWFHRHROEAFR 480
QY 1088 TTCTCAATGCTCTT 1101
DB 481 EALHMDRYLLHPD 494

```

```

RESULT 9
US-10-141-756-462

```

```

; Sequence 462, Application US/10141756
; Publication No. US20030207359A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen

```

```

; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C200
; CURRENT APPLICATION NUMBER: US/10/141,756
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 462
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-756-462

```

```

Query Match      2.6%; Score 36.4; DB 17; Length 600;
Best Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;

```

```

QY 608 AGACCTACATGATGCGCATCTTAAAGCAGGATCTGAGACATTTCTAAGCATCCATA 667
DB 1 MRSCLMRCRHLSQGVMSLLAVLFLPALPSPFKEPQTRSRHQRTENIKERSLSLSIA 60
QY 668 TAGGATGAGCGCGTAAAGTCCGATTGAAGAGATGAAGGGGTGTGAAGGTGTGTC 727
DB 61 KKSQAPPRARRTTTYEAPENNALNTOTQPKATTTGDRKEANQAPPEBDKVPHTAQ 120
QY 728 AAAAGAGGTCGATGGCTATATACAGCGCTAAGCAGGTGGCTAGCAGCTGTGCAGC 787
DB 121 RAAMKSPKEKTMVNTLSPRGDAGMASGRTEASWSQDRTKTQNGGQTRKLTASRTV 180
QY 788 TGTGAATAAGTCATCTGTTAGTATGTCACCACTTAATGTGACGAGTGAATGCTGAT 847
DB 181 SEKHQKATTAATKTLIPKSQHRMLAPTGAVSTRROKGVTTAVIPPEKKEPPATPPAPF 240
QY 848 TGGGTTAAATGGCATGTAGTGTAGTCCGAAAAACAGTTAGATCTAGTTAAAGGA 907
DB 241 QSPFTQNRORLKAANFKSEPRMDPEBKYSFEIGLQTTCPDSVKIKASISLMLOKLFYN 300
QY 908 AGCTGAAGCTGAACCTGTGAGAAATAAGCCTGTTGAATACAGCTTGATTAACCAATT 967
DB 301 LTLFLDSHRHFNQSEWDRLEHFAAPPGFMELNYSLVQKVTRPPVPOQQLLASLPAGSL 360
QY 968 CAGTGTCAAGGCTGCTCGTATATGCTGAGCTTCCCTGTGCAATTTGGGGTAACTATT 1027
DB 361 RCITCAVVGNGILNNSHMGDEIDSHDVYFRLSGALIKGYEODVGTSTRSFYFTAFSLQ 420
QY 1028 TCATAGTGGGAGAGATGCAACTCTATTTCATTTGAATGAATCTAAGCATCTCGGAGAG 1087
DB 421 SLLILGNRFKNVPLGKDVRYLHLEGRDYEMLEALMNQTVMSKNLFWFRHRPOEAFR 480
QY 1088 TTCTCAATGCTCTT 1101
DB 481 EALHMDRYLLHPD 494

```

```

RESULT 10
US-10-141-759-462
; Sequence 462, Application US/10141759
; Publication No. US20030207361A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baretini, Maureen

```

```

; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C197
; CURRENT APPLICATION NUMBER: US/10/141,759
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 462
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-759-462

```

```

Query Match      2.6%; Score 36.4; DB 17; Length 600;
Best Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;

```

```

QY 608 AGACCTACATGATGCGCATCTTAAAGCAGGATCTGAGACATTTCTAAGCATCCATA 667
DB 1 MRSCLMRCRHLSQGVMSLLAVLFLPALPSPFKEPQTRSRHQRTENIKERSLSLSIA 60
QY 668 TAGGATGAGCGCGTAAAGTCCGATTGAAGAGATGAAGGGGTGTGAAGGTGTGTC 727
DB 61 KKSQAPPRARRTTTYEAPENNALNTOTQPKATTTGDRKEANQAPPEBDKVPHTAQ 120
QY 728 AAAAGAGGTCGATGGCTATATACAGCGCTAAGCAGGTGGCTAGCAGCTGTGCAGC 787
DB 121 RAAMKSPKEKTMVNTLSPRGDAGMASGRTEASWSQDRTKTQNGGQTRKLTASRTV 180
QY 788 TGTGAATAAGTCATCTGTTAGTATGTCACCACTTAATGTGACGAGTGAATGCTGAT 847
DB 181 SEKHQKATTAATKTLIPKSQHRMLAPTGAVSTRROKGVTTAVIPPEKKEPPATPPAPF 240
QY 848 TGGGTTAAATGGCATGTAGTGTAGTCCGAAAAACAGTTAGATCTAGTTAAAGGA 907
DB 241 QSPFTQNRORLKAANFKSEPRMDPEBKYSFEIGLQTTCPDSVKIKASISLMLOKLFYN 300
QY 908 AGCTGAAGCTGAACCTGTGAGAAATAAGCCTGTTGAATACAGCTTGATTAACCAATT 967
DB 301 LTLFLDSHRHFNQSEWDRLEHFAAPPGFMELNYSLVQKVTRPPVPOQQLLASLPAGSL 360
QY 968 CAGTGTCAAGGCTGCTCGTATATGCTGAGCTTCCCTGTGCAATTTGGGGTAACTATT 1027
DB 361 RCITCAVVGNGILNNSHMGDEIDSHDVYFRLSGALIKGYEODVGTSTRSFYFTAFSLQ 420
QY 1028 TCATAGTGGGAGAGATGCAACTCTATTTCATTTGAATGAATCTAAGCATCTCGGAGAG 1087
DB 421 SLLILGNRFKNVPLGKDVRYLHLEGRDYEMLEALMNQTVMSKNLFWFRHRPOEAFR 480
QY 1088 TTCTCAATGCTCTT 1101
DB 481 EALHMDRYLLHPD 494

```

```

RESULT 11
US-10-140-805-462
; Sequence 462, Application US/10140805
; Publication No. US20030207417A1

```

```
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: DeForge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ TITLE OF INVENTION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: P3330R1C176
/ CURRENT APPLICATION NUMBER: US/10/140,805
/ CURRENT FILING DATE: 2002-05-07
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 462
/ LENGTH: 600
/ TYPE: PRP
/ ORGANISM: Homo Sapien
US-10-140-805-462
```

```
Query Match      2.6%; Score 36.4; DB 17; Length 600;
Best Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;
```

```
QY      608 AGACCTACATGATGCCATCTTAAGAGGAGATCTGAGACATTTTCTAAGCATCCATA 667
      1 MNSCLMRCHLSQGVMSLLAVLFPLALPSFIKEPQTKSRHQRTENIKERSLQSLA 60
QY      668 TAGGCATTGGGGCGGCGTAAGTCCGCGATTGAAGAGATAGGGGGGTGGAAGTGTGTGTC 727
      61 KKSQAPTRARRTTYABRAPENNALNTOTPRKATTTGDRKBNAPREEDKVPHTAQ 120
      728 AAAAGAGGTCGATTGGCTATACACGCCGCTAAGCAGGCGGTAGCGACTGTCCAGC 787
      121 RAAMKSPKEKTMVNTLSPRGODAGMASRTEAQSMSODTKTQNGSQTRKLTASRTV 180
QY      788 TGTGAATAAGCTACTTGTAGTATGTCCACCTAATGTACAGAGATGCAATGCTGAT 847
      181 SKHQKAKATTAKTLIPKSOHMLAPTAGVSTRTRQKGVTTAVIPPEKEKPPQATPPAPF 240
QY      848 TGGGTTAAATGGGCGATGTAGTGTAGGTGCCGAAAACAGTTAGTCTAGTTAAAGGA 907
      241 QSPFTTORNRQLKAANFKSEPRWDPEEKYSFELGSLQTTCPDSVKIKASLMLQKLFPLN 300
QY      908 AGCTGAAAGCTGAACCTGTCAAGAAATAAGCCTTTGGAATACAGCTGATTAACCAATT 967
      301 LTLFLDSHFHNSQSEWDLRHFAPRPFGEMLNYSLVQKVTRPPVPOOQLLASLPAGSL 360
QY      968 CAGTGTCAAGGCGTGTCTGATATGCTGAGACTTCCCTGCTGCATTTGGGGTAATATT 1027
      361 RCITCAVVGNGGILNNSHMGGEIDSHDYVFRLSGALIKGYEODVGTRTSFYGTAFSLQ 420
QY      1028 TCATAGTGGGCGAGATGCAACTCTATTCTTCAATTGAATCTAATCTATTCTGGGTAGG 1087
      421 SLLILNNGFKNVPYLGKQVRYLHFLEGRDYEWLEALLMNQVMSKNLFWFHRHROEAFR 480
QY      1088 TTCTCAATGCTCTT 1101
      481 EALHMDRYLLLHPD 494
DB
```

RESULT 12

```
US-10-140-864-462
/ Sequence 462, Application US/10140864
/ Publication No. US20030207419A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: DeForge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ TITLE OF INVENTION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: P3330R1C184
/ CURRENT APPLICATION NUMBER: US/10/140,864
/ CURRENT FILING DATE: 2002-05-07
/ Prior Application removed - See Palm or File Wrapper
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 462
/ LENGTH: 600
/ TYPE: PRP
/ ORGANISM: Homo Sapien
US-10-140-864-462
```

```
Query Match      2.6%; Score 36.4; DB 17; Length 600;
Best Local Similarity 9.7%; Pred. No. 1.5;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;
```

```
QY      608 AGACCTACATGATGCCATCTTAAGAGGAGATCTGAGACATTTTCTAAGCATCCATA 667
      1 MNSCLMRCHLSQGVMSLLAVLFPLALPSFIKEPQTKSRHQRTENIKERSLQSLA 60
QY      668 TAGGCATTGGGGCGGCGTAAGTCCGCGATTGAAGAGATAGGGGGGTGGAAGTGTGTGTC 727
      61 KKSQAPTRARRTTYABRAPENNALNTOTPRKATTTGDRKBNAPREEDKVPHTAQ 120
      728 AAAAGAGGTCGATTGGCTATACACGCCGCTAAGCAGGCGGTAGCGACTGTCCAGC 787
      121 RAAMKSPKEKTMVNTLSPRGODAGMASRTEAQSMSODTKTQNGSQTRKLTASRTV 180
QY      788 TGTGAATAAGCTACTTGTAGTATGTCCACCTAATGTACAGAGATGCAATGCTGAT 847
      181 SKHQKAKATTAKTLIPKSOHMLAPTAGVSTRTRQKGVTTAVIPPEKEKPPQATPPAPF 240
QY      848 TGGGTTAAATGGGCGATGTAGTGTAGGTGCCGAAAACAGTTAGTCTAGTTAAAGGA 907
      241 QSPFTTORNRQLKAANFKSEPRWDPEEKYSFELGSLQTTCPDSVKIKASLMLQKLFPLN 300
QY      908 AGCTGAAAGCTGAACCTGTCAAGAAATAAGCCTTTGGAATACAGCTGATTAACCAATT 967
      301 LTLFLDSHFHNSQSEWDLRHFAPRPFGEMLNYSLVQKVTRPPVPOOQLLASLPAGSL 360
QY      968 CAGTGTCAAGGCGTGTCTGATATGCTGAGACTTCCCTGCTGCATTTGGGGTAATATT 1027
      361 RCITCAVVGNGGILNNSHMGGEIDSHDYVFRLSGALIKGYEODVGTRTSFYGTAFSLQ 420
QY      1028 TCATAGTGGGCGAGATGCAACTCTATTCTTCAATTGAATCTAATCTATTCTGGGTAGG 1087
      421 SLLILNNGFKNVPYLGKQVRYLHFLEGRDYEWLEALLMNQVMSKNLFWFHRHROEAFR 480
QY      1088 TTCTCAATGCTCTT 1101
      481 EALHMDRYLLLHPD 494
DB
```

```
RESULT 13
US-10-142-426-462
; Sequence 462, Application US/10142426
; Publication No. US20040048333A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P333081C224
; CURRENT APPLICATION NUMBER: US/10/142,426
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 462
; LENGTH: 600
; TYPE: PRP
; ORGANISM: Homo Sapien
US-10-142-426-462

Query Match
Best Local Similarity 2.6%; Score 36.4; DB 17; Length 600;
Matches 48; Conservative 125; Mismatches 321; Indels 0; Gaps 0;

QY 608 AACCTACATGATGCCACATCTTAAAGCAGGATCGACATTTTCTAAGCATCCATA 667
DB 1 MSCCMLRCHLSQGVWMSLLAVLFPLPSPFIKBPQTKPSRIQRTENITERSLSQSLA 60
QY 668 TAGGATTTGGGCGCTAAGCGGCGCATTAAGAGAGATGAAGGCGGTGAAGTGTGTCTC 727
DB 61 KKSQAPTRARRTTYIABEPAPENMALNTQTPKATHTTGRKEANQAPREBDKVPHTAQ 120
QY 728 AAAAGAGGTGCTATGGCTATACCAAGCGGCTAAGCAGGTGGGCTAAGCAGCTGTGCACG 787
DB 121 RAAMKSPKEXTMVTLSRPGDAGMASGRTEAQSWKSODTKTQONGGQTKRLTFASRTV 180
QY 788 TGTGAATTAAGCTCAGTCTTGAAGTATGTCACCTTAATGTCAGCAGATCAATGTGTAT 847
DB 181 SEKHOGKATTTAKTLIPKSOHRMLAPRTGAVSTRTRQKGVTTTIVIPKEKKPQATPPAPF 240
QY 848 TGGGTTAAATAGGAGCATGATGATGAGTCCGAAAAACAGCTTAAGTCTAAGTTAAAGGGA 907
DB 241 QSPPTQORNLKAAAFKSPRWDPEEKISFEIGLQTTTPDSVKIKASSLMLQKLPFN 300
QY 908 AACTGAAAGCTGAACCTGTCAAGAAATTAAGCTTTGGAATACCAAGCTTGAATACCAANT 967
DB 301 LTLPLDSRRHFNQSEMDRLHFAAPRFEMELNYSLVQKVTTRPPVYQQQLLASLPAAGSL 360
QY 968 CAGTGTCTCAGGCTGCTCGATATGCTGAGCTTCCCTGTGCGATTTGGGGTAAGTACTT 1027
DB 361 RCITCAVAVNGGILNNSHMGQEIASHDYVFRLSGALIKINGEYEDVGTRTSYFTAFSLTQ 420
QY 1028 TCATAGTGGGAGATGCAACTATTTTCAATTGAATGAACCTAAACATTTCTGGGTAGAG 1087
DB 421 SLITIGNRGFPKVPKIGKDVRYLHPLEGTQDYEWLEALLMNQTVMSKNLFWFHRHPOEAFR 480
```

```
QY 1088 TTCTCATGCTCTT 1101
DB 481 EALHMDRYLLHPD 494

RESULT 14
US-09-775-938A-34/C
; Sequence 34, Application US/09775938A
; Patent No. US20020081665A1
; GENERAL INFORMATION:
; APPLICANT: Haygood, M.
; APPLICANT: Davidson, S.K.
; APPLICANT: Allen, S.W.
; APPLICANT: Hildebrand, M.
; TITLE OF INVENTION: Bryostatins, Bryopyrans and Polyketides: Compositions and Methods
; FILE REFERENCE: 1133.010US1
; CURRENT APPLICATION NUMBER: US/09/775,938A
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: PCT/US00/21326
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/147,283
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 2672
; TYPE: DNA
; ORGANISM: Endobugula sertula
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2672)
; OTHER INFORMATION: N refers to any nucleotide.
US-09-775-938A-34

Query Match
Best Local Similarity 2.6%; Score 36.2; DB 9; Length 2672;
Matches 41; Conservative 60; Mismatches 65; Indels 0; Gaps 0;

QY 674 TTGGCGGTGAAGTCGCGATTAAGAGATTAAGGGGCTGTGAAGTGTGTCTCAAAAG 733
DB 702 TTCMGCMGRSTCAKSWYKTKKWTWCCAAAGGSTRKSKGTWTWYSGWMANCRG 643
QY 734 AGTGATTTGGCTATACCAAGCGGCTAAAGCAGTGGGCTAGACGCTGTGACGCTGTGA 793
DB 642 MYSWGTTTTCRAAWKSSGYTKCGTGTAGTGYWCCAWWMSCRMYSGRSGSKCTTKWT 583
QY 794 TAACGCTAGTGTAGGTATGTCCACTAATGTACAGAGATGCAAAATGC 843
DB 582 KAAWAMMCTYTTTGSAAWAASTYTKWAGGKWKCGKCKMCRAMTKY 533

RESULT 15
US-10-312-841-1/C
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; PRIOR FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 367378
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1
```


Query Match 2.6%; Score 36; DB 16; Length 3673778;
 Best Local Similarity 56.9%; Pred. No. 3.3e+02;
 Matches 66; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
 QY 1243 AAAAGAGCTTGAGAGACCATTCAAATCCTAAGGCTCTCTTTCTTTCTGCAATCAATCA 1302
 Db 604282 AAAATTTACATTAATTAATCATCTTAATCATATCTCCAAACTTATATCTTTCCCGCTA 604223
 QY 1303 AGAATCATACACTCAAACGAGACTCTTTCTATCTTCCCTATAGCAATTCCTCAA 1358
 Db 604222 ACAATTAATACGTTACAACTCTTAATCTATCATATATATACAAATAATATCTAA 604167

Search completed: February 9, 2005, 15:26:37
 Job time : 769 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2005, 10:10:25 ; Search time 4286 Seconds
(without alignments)

12238.116 Million cell updates/sec

Title: US-10-070-386-1

Perfect score: 1378
Sequence: 1 gtgcagctg99tgcgtacatc.....accatcaatcaactaaca 1378

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41.6	3.0	557	BH285801	CH230-118
2	39	2.8	1031	CNS0500M	ALJ39601 Tetradon
3	38.8	2.8	1205	CC264185	CC264185 CH261-57E
4	38.4	2.8	673	BZ246183	BZ246183 CH230-374
5	38.4	2.8	734	AG590975	AG590975 Mus muscu
6	38.2	2.8	923	AO018895	AO018895 CIT-HSP-2
7	38.2	2.8	479	CNS0489T	AL299018 Tetradon
8	38	2.8	482	BG185531	BG185531 RST4479 A
9	37.8	2.7	1101	CNS0000D1	AL065414 Drosophila
10	37.6	2.7	688	AV384223	AV384223 AV384223
11	37.4	2.7	598	AZ960759	AZ960759 2M0228L20
12	37.2	2.7	250	CE251889	CE251889 tigr-gsa-
13	37.2	2.7	625	CN084563	CN084563 EC2BBA24B
14	37	2.7	446	CL700776	CL700776 SP_Ba006
15	37	2.7	590	CE571358	CE571358 tigr-gsa-
16	37	2.7	720	CL705705	CL705705 SP_Bb004
17	37	2.7	625	CE686685	CE686685 tigr-gsa-
18	37	2.7	801	BG534705	BG534705 602553868
19	36.8	2.7	562	CE834892	CE834892 tigr-gsa-
20	36.8	2.7	704	CL197184	CL197184 104_424_1
21	36.8	2.7	730	CN975554	CN975554 25417_125
22	36.8	2.7	738	CL157592	CL157592 104_345_1
23	36.8	2.7	773	CL157591	CL157591 104_345_1
24	36.8	2.7	1101	CNS016K9	AL106851 Drosophila

C 25	36.6	2.7	1075	8	CC280186	CC280186 CH261-112
26	36.4	2.6	778	7	CV490143	CV490143 AGENCOURT
27	36.4	2.6	1030	5	BX460908	BX460908 BX460908
28	36.2	2.6	457	4	BI741102	BI741102 GC93511.Y
29	36.2	2.6	634	4	AZ037814	AZ037814 RPT-23-2
30	36	2.6	532	1	AL923574	AL923574 AL923574
31	36	2.6	551	8	AZ163313	AZ163313 SP_0073.B
32	36	2.6	621	5	BP504692	BP504692 BP504692
33	36	2.6	683	1	AA556936	AA556936 778 Lobi
34	36	2.6	1201	9	CNS0102D	AL098431 Drosophila
35	35.8	2.6	432	4	BI502064	BI502064 tm08105.Y
36	35.8	2.6	474	4	BI501745	BI501745 tm04605.Y
37	35.8	2.6	592	8	AZ838544	AZ838544 2M0134H15
38	35.8	2.6	610	8	AO891964	AO891964 HS_3142.A
39	35.8	2.6	845	8	BZ114424	BZ114424 CH230-510
40	35.8	2.6	891	6	CD789126	CD789126 EST660487
41	35.8	2.6	945	6	CD792880	CD792880 EST664241
42	35.8	2.6	1306	8	BZ565154	BZ565154 PAC82-164
43	35.6	2.6	644	8	BH766160	BH766160 BMBAC359H
44	35.6	2.6	748	4	BI886660	BI886660 ZF637-1.0
45	35.6	2.6	814	8	BH109031	BH109031 RPT-24-3

ALIGNMENTS

RESULT 1
BH285801
LOCUS BH285801 557 bp DNA linear GSS 30-NOV-2001
DEFINITION CH230-118D10.TVB CHORI-230 Segment 1 Rattus norvegicus genomic
clone CH230-118D10, genomic survey sequence.
BH285801
ACCESSION BH285801
VERSION BH285801.1 GI:17198207
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 557)
AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 EORI segment
Unpublished (1999)
JOURNAL Other_GSSs: CH230-118D10.TVB
COMMENT Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pierre de Jong (pjejong@mail.cho.org).
Clones may be purchased from BACPC Resources
(http://www.chori.org/bacpac/or eting_information.htm). BAC end
page: http://www.tigr.org/bacpac/or eting_information.htm1
Plate: 118 row: D column: 10
Seq primer: 17
Class: BAC ends.

FEATURES

source
Location/Qualifiers
1..557
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SaNHsd/MCM"
/db_xref="taxon:10115"
/clone="CH230-118D10"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/note="Vector: pTARBAC2.1; Site_1: EORI; Site_2: EORI;

CHORI-230 Rat (BN/SnHsd/MCM) BAC library produced by
Pleeter de Jong"

Query Match 3.0%; Score 41.6; DB 8; Length 557;
Best Local Similarity 49.1%; Pred. No. 0.66; Mismatches 114; Indels 0; Gaps 0;
Matches 110; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

ORIGIN
Qy 1155 AGAGTGGCGCATTTAGTAGTACCGCATCGAACCCGGAGGGGTTCAAGACATGGGCGT 1214
|||
Db 89 AGATCCCTGGAATTAAGAGATGACAAAGAGAGGGGAGACATATGAAGATCATCAGG 148
Qy 1215 ACGTGATACATAGAGTCAATAGAAACATTAAGAGGCTTGAAGAACCATTTCAATCTTA 1274
|||
Db 149 ACATGTATGACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 208
Qy 1275 GGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1334
|||
Db 209 CGTTATCATGTTTCTTTCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 268
Qy 1335 ATCTTCCCTATAGCAATTCCTCAAAACCATCAATCACTTACCA 1378
|||
Db 269 GACTGACTAGAGGAATCCCAACCTCTAGATGATACATGACA 312

RESULT 2
CNS050QW/c 1031 bp DNA linear GSS 01-SEP-2000
LOCUS Tetradon nigroviridis genome survey sequence Sp6 end of clone
DEFINITION 025803 of library B from Tetradon nigroviridis, genomic survey
sequence.

ACCESSION AL349601 GI:8243371
VERSION AL349601.1
KEYWORDS GSS; genome survey sequence.
SOURCE Tetradon nigroviridis
ORGANISM Tetradon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorphi; Acanthopterygii; Percomorphi; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE 1
AUTHORS Roest Crolius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
Saurin, W., and Weissenbach, J.
TITLE Estimate of human gene number provided by genome-wide analysis
using Tetradon nigroviridis DNA sequence
JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE 20296633
PUBMED 10835645

REFERENCE 2
AUTHORS Roest Crolius, H., Jallion, O., Dasilva, C., Ozouf-Costaz, C.,
Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,
Saurin, W., Bernot, A., and Weissenbach, J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish tetradon nigroviridis
JOURNAL Genome Res. 10 (7), 939-949 (2000)
MEDLINE 10899143
PUBMED 10899143

REFERENCE 3
AUTHORS (bases 1 to 1031)
TITLE Genoscope.
DIRECT SUBMISSION Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segreff@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
- Note="Vector: pTARBAC2.1; Site 1: EcorI; Site 2: EcorI;
CH261 Female Chicken library - For library and clone
ordering information: http://www.chori.org/bacpac"

FEATURES
source 1..1031
/organism="Tetradon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:9983"

/clone="025B03"
/clone_lib="B"
/note="Genoscope sequence ID : COAB025CA02B1-end : Sp6"

ORIGIN
Query Match 2.8%; Score 39; DB 9; Length 1031;
Best Local Similarity 43.1%; Pred. No. 5.3; Mismatches 107; Indels 0; Gaps 0;
Matches 81; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

ORIGIN
Qy 792 AATAAGCTCACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 851
|||
Db 523 AAAAAAGTTTGTTCGTAGACAGCTNACTAATAAATTAACNAGAAANAGAAATTAAGT 464
Qy 852 TTAATAAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 911
|||
Db 463 TTNAAATGAGNANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 404
Qy 912 GAAAGCTGACCTGTCGAAATTAAGCCGCTGGAATCAAGGTTGATACCAATTCAGT 971
|||
Db 403 NAATNCTTACCATATNNAAAAACGNANNTATGNNAACTNAGTTCAAAATCCNACTCTTT 344
Qy 972 CGTCAGG 979
|||
Db 343 CTTNAGG 336

RESULT 3
CC264185 1205 bp DNA linear GSS 13-MAY-2003
LOCUS CH261-57E7_Sp6.1 CH261 Gallus gallus genomic clone CH261-57E7,
DEFINITION genomic survey sequence.

ACCESSION CC264185 GI:30610833
VERSION CC264185
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianidae; Gallus.

REFERENCE 1
AUTHORS 1 (bases 1 to 1205)
Kremetzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.
TITLE Gallus gallus BAC End Reads
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: Sp6 ATTTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 50
High quality sequence stop: 729.

FEATURES
source 1..1205
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-57E7"
/sex="female"
/cell_line="UCDD001, inbred 256"
/clone_lib="CH261"
/note="Vector: pTARBAC2.1; Site 1: EcorI; Site 2: EcorI;
CH261 Female Chicken library - For library and clone
ordering information: http://www.chori.org/bacpac"

ORIGIN
Query Match 2.8%; Score 38.8; DB 8; Length 1205;
Best Local Similarity 48.2%; Pred. No. 6.4; Mismatches 117; Indels 0; Gaps 0;
Matches 109; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

Qy 823 AATGTACAGATGCAAAATGCTGATTTGGTTTAAATGGGATGATGATGATGATGATGATGAT 882

Db	606	TTCTGTTTAAATTAATAGAAAATGCTATTATATATATATATATATACACACAGAGA	547
Qy	1160	GGCGCATTTGAATGCTACCGCATCGAATCCCGGAAGCGGTTCAAGACATGGCGCTACTGA	1219
Db	546	GAGGAGACAGACAGACAGACAGACACAGACACAGACACAGACACAGACACAGACA	487
Qy	1220	GATACATAGAGTCAATAGAAAATATAAAGAGGCTTGAAAGAACCATTCCTTAATCTTAAG	1275
Db	486	CACAGATTACAGACACACACACAGAGGAGATTCTGCAGAGCTTCTCAAGAGAAAG	431
RESULT 5	AG590975/c		
LOCUS	AG590975	734 bp	linear
DEFINITION	Mus musculus molossinus DNA, clone:MSMg01-519M14.T7, genomic survey		
ACCESSION	AG590975		
VERSION	AG590975.1	GI:48351805	
KEYWORDS	GSS.		
SOURCE	Mus musculus molossinus		
ORGANISM	Mus musculus molossinus		
REFERENCE	1	Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.	
REFERENCE	2	BAC end Sequences of Library MSMg01	
REFERENCE	3	Unpublished	
REFERENCE	4	(bases 1 to 734)	
REFERENCE	5	Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.	
REFERENCE	6	Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical	
REFERENCE	7	and Chemical Research (RIKEN), Genomic Sciences Center (GSC),	
REFERENCE	8	1-7-22 Suehiro-chou, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan	
REFERENCE	9	(e-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,	
REFERENCE	10	Tel:81-45-503-9111, Fax:81-45-503-9170)	
REFERENCE	11	Clones are derived from the mouse BAC library MSMg01. For BAC	
REFERENCE	12	library availability, please contact Kuniya Abe (abe@rtc.riken.jp).	
REFERENCE	13	Tsukuba Institute, Bio Resource Center.	
REFERENCE	14	The Institute of Physical and Chemical Research (RIKEN) 3-1-1	
REFERENCE	15	Koyadai, Tsukuba, 305-0074 Japan	
REFERENCE	16	phone: 81-298-36-9189, fax: 81-298-36-9199	
REFERENCE	17	e-mail: abe@rtc.riken.jp	
REFERENCE	18	PRIMERS	
REFERENCE	19	Sequencing : T7	
REFERENCE	20	LIBRARY	
REFERENCE	21	Vector : pBACe3.6	
REFERENCE	22	R.Site 1 : EcoRI	
REFERENCE	23	R.Site 2 : EcoRI	
REFERENCE	24	Location/Qualifiers	
REFERENCE	25	1. 734	
REFERENCE	26	/organism="Mus musculus molossinus"	
REFERENCE	27	/mol_type="genomic DNA"	
REFERENCE	28	/sub_species="molossinus"	
REFERENCE	29	/db_xref="taxon:57486"	
REFERENCE	30	/clone="MSMg01-519M14.T7"	
REFERENCE	31	/sex="male"	
REFERENCE	32	/tissue_type="mixture of kidney and spleen"	
REFERENCE	33	/clone_lib="MSMg01 Mouse Male BAC Library"	
ORIGIN			
Query Match	2.8%	Score 38.4	DB 9; Length 734;
Best Local Similarity	54.2%	Pred. No. 7.1;	
Matches	78; Conservative	0; Mismatches	66; Indels 0; Gaps 0;
Qy	1106	CTGTCACTTACACACATCATGGGGGCTCAAAAGCTATACAGCTTCAATAGAGAGCGCGCA	1165
Db	670	CAGACAGACACACACAGCAAGAGAGACAGAAACAGACACAGAAAGAGAGGACACAGA	611
Qy	1166	TTGAAGTAGCTACCGCATCGAATCCCGGAACGGGTTCAAGACATGGCGGCTAGATATACA	1225
Db	610	GACACACAGATTAGAGAGACACACACAGAAATTAAGAGAGAGACAGAGACAGAGGACAGACA	551
Qy	1226	TAGAGTCATAGAAAATTAAGAAAGG	1249

Db 550 GAGAGATACAGATACAGACAGAGA 527

|||||

RESULT 6
LOCUS A0018895/c 479 bp DNA linear GSS 09-JUN-1998
DEFINITION CIT-HSP-2302116.TF CIT-HSP Homo sapiens genomic clone 2302116,
genomic survey sequence.
ACCESSION A0018895
VERSION A0018895
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 479)
Adams,M.D., Rounley,S.D., Zhao,S., Field,C.E., Baas,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
Unpublished (1998)
JOURNAL Other_GSSs: CIT-HSP-2302116.TR
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamas@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/cdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Classes: BAC ends.

FEATURES
Location/Qualifiers
source 1..479
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2302116"
/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/note="Vector: pBelBAC11; Site_1: HindIII; Site_2:
HindIII"

ORIGIN
Query Match 2.8%; Score 38.2; DB 8; Length 479;
Best Local Similarity 50.3%; Pred. No. 7;
Matches 94; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 31 GGTGTCGAAGTCTGTAGTACATGAGATGAGAAAGCGTGTGTCGACGT 90
|||
Db 460 GGTGTCGAAGTCTGTAGTACATGAGATGAGAAAGCGTGTGTCGACGT 401
|||
QY 91 GAAGCTGAAGAGAGAGCAAGATTAATCACAATCGATCGTTGATCATGCTT 150
|||
Db 400 AGTGAAGAAATGAGAGCGTATGATATATGAAATGTTATATCGCTTGGC 341
|||
QY 151 TCAAGAGACACGTTGATCTACCTGGTGTCCCTCTTCGTTGTAACAAGATCAATTC 210
|||
Db 340 TTGGCAATGACATTTGTAAGATTCCTCTCTTTGTTGAAATGTTCAATCAT 281
|||
QY 211 GGATGAC 217
|||
Db 280 GGAAAC 274
|||

RESULT 7
LOCUS CNS04NPT/c 923 bp DNA linear GSS 01-SEP-2000

DEFINITION Tetradon nigroviridis genome survey sequence FUC-Or1 end of clone
123C09 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL299018
VERSION AL299018.1 GI:8038159
KEYWORDS GSS: genome survey sequence.
SOURCE Tetradon nigroviridis
ORGANISM Tetradon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.

REFERENCE 1
Roest Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F.,
Saurin,M. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetradon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)

JOURNAL MEDLINE
PUBMED 20296633
REFERENCE 10835645
2

TITLE
Roest Crolius,H., Jallion,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,M., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradon nigroviridis
Genome Res. 10 (7), 939-949 (2000)

JOURNAL MEDLINE
PUBMED 10899143
REFERENCE 3 (bases 1 to 923)
Genoscope.
Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetradon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetradon.

FEATURES
Location/Qualifiers
source 1..923
/organism="Tetradon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="123C09"
/clone_lib="G"
/note="Genoscope sequence ID : COB6123AB05SP1-end :
FUC-Or1"

ORIGIN
Query Match 2.8%; Score 38.2; DB 9; Length 923;
Best Local Similarity 50.3%; Pred. No. 8; 9;
Matches 85; Conservative 3; Mismatches 81; Indels 0; Gaps 0;

QY 891 AGATCTAGTTAAAGGAAGCTGAAGCTGAACCTGTCAAGAAATAAGCTGTTGAATACA 950
|||
Db 352 RGAGAAAGAGAAAAAGCCATGCTGTGTCTATCAACAATATGTGTTGGCTCG 293
|||
QY 951 ACGTTGATTAACCAATTCAGTCGTCAAGGGTGTCTGTATATGCTGACCTCCCTGTGGC 1010
|||
Db 292 GCGTTAATATCTTAATAATGTTGGATGAAAGCCGACTGGCGGCTTTCGCTTAAAT 233
|||
QY 1011 ATTGGGGTAACTATTTCATAGTGGGCGAGAAATGCACTGATTTTCA 1059
|||
Db 232 ATCTCCGCTCCCTGTTGATTCGCGGTGAACAAAGACAAATATCA 184
|||

RESULT 8
LOCUS BG185531/c 482 bp mRNA linear EST 21-APR-2001
DEFINITION R814479 Athysys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG185531

VERSION	EST.	GI:13707218
KEYWORDS	BJ185531.1	GI:13707218
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. (bases 1 to 482)	
AUTHORS	Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramchandran,R., Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith,E., Veloso,N., Kilka,A., Hees,J., Cothren,K., Lo,K., Offenbacher,J., Danzig,J. and Ducar,M.	
TITLE	Creation of genome-wide protein expression libraries using random activation of gene-wide protein expression	
JOURNAL	Nat. Biotechnol. 19 (5), 440-445 (2001)	
MEDLINE	21227151	
PUBMED	11329013	
COMMENT	Contact: Scott J. Cain Athersys, Inc. 3201 Carnegie Ave, Cleveland, OH 44115, USA Tel: 216 431 9900 Fax: 216 361 9536 Email: scain@athersys.com High quality sequence stop: 451. Location/Qualifiers 1..482	
FEATURES	Source /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /cell_line="HT1080" /clone_lib="Athersys RAGE library" /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology', in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."	
ORIGIN	Query Match Best Local Similarity 2.8%; Score 38; DB 4; Length 482; Matches 71; Conservative 0; Mismatches 57; Indels 0; Gaps 0;	
QY	913 AAAGCTGAACCTGTCAGAAATAGCCTTTGGAATACACCTGTGTAACCAATTCAGTC	972
DB	164 AGAGGCACCCCTTTCTGTACAGACATTTGAATATCATGCTATCTCATTTGGAT	105
QY	973 GTCAAGGCTGCTGTGATATGCTGAGCTTCCTGTGCGATTGGGGTAACTATTTCATA	1032
DB	104 TTCAAGTGTTCCTTCCTCCCACTCNNAACGCGATTCGGAGTTTGGAAATGTTGGAA	45
QY	1033 GTGGGGCA	1040
DB	44 GTATTGNA	37
RESULT 9		
CNS000D1/c		
LOCUS	CNS000D1	1101 bp DNA linear GSS 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR01J16 of RPIC-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	
ACCESSION	AL065414	
VERSION	AL065414.1	GI:4938827
KEYWORDS	GSS.	
SOURCE	Drosophila melanogaster (fruit fly)	
ORGANISM	Drosophila melanogaster	
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
AUTHORS	1 (bases 1 to 1101)	
TITLE	Genoscope. Direct Submission	

JOURNAL

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqret@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Ito Oseegawa and Aaron Mammone in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Source	Location/Qualifiers
	1..1101
	/organism="Drosophila melanogaster"
	/mol_type="genomic DNA"
	/db_xref="taxon:7227"
	/clone="BACR01J16"
	/clone_lib="RPCI-98"
	/note="end : TEST"

ORIGIN

Query Match	2.7%; Score 37.8; DB 9; Length 1101;
Best Local Similarity	15.5%; Pred. No.13; Indels 1; Gaps 1;
Matches	51; Conservative 155; Mismatches 122;

```

QY      1049  CTCATATTTCATGATCATTAATACTATTCTGGGTAGAGTTCCATGCGTCTCGTG 1108
          ||::| | ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       1099  MTNNHNYMTNMNHNHTNTMNTNMNTMTMTWTTHNTTNNTNMNTTNMTMTGTCMY 1040C
QY      1109  TCACTTAACAATCATGAGGGGTCAACAACGATATACAGTTCATGAAGAGTGGGCATTG 1168
          ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       1039  MCNNHHNMAMMMCMYTWTYYTMNMCMNYHCSTYYYUUYUUTAAAHYMAHNSHA 980
QY      1169  AAGTAGTACCAGTCAACCCGGAACGGTTCAGACATGGGCGTACGATATACATAG 1228B
          |:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       979  TMTNNHYNMCMNMNHNCSTCYTACMYCMYCMYMAWMAWMAACMYMNCMTMNHN 920
QY      1229  AGTCATGAACAATGAAGAAGAGGTGAAGAASCATTAATATCCTAAGAGTCTCTCTT 1288B
          ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       919  WCATCTCTAATNNHAMNNHTANNNHMNNHTCTGYUYNHYUYNHYUYNHYUWT-TW 861
QY      1289  TTGCAATCATCAAGATCATACACTAACACAGAACTCTTTCTATCTTCCSTATAGC 1348B
          ||::| | ::| | ::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       860  TMNHTAHYMAAYAMMYUUYCTMMNHTNMHTCYAHNAHWYTYUTYUYNHNCSTCGCY 801
QY      1349  AATTTCCAAAACCATCATCAACCTTAC 1377
          ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       800  HCHCCSMHYMNCAMMMCMNMHNCYTCA Y 772
    
```

RESULT 10

AV384223	688 bp	mRNA	linear	EST 27-OCT-1999
LOCUS	AV384223			
DEFINITION	AV384223 Halocynthia roretzi Fertilized egg			
ACCESSION	CNRA clone 008M08_5' , mRNA sequence.			
VERSION	AV384223			
KEYWORDS	AV384223.1 GI:6129280			
SOURCE	EST.			
ORGANISM	Halocynthia roretzi			
REFERENCE	Hakabe,K.W.			
TITLE	Halocynthia roretzi EST			
JOURNAL	Unpublished (1999)			

QY 329 AACGAAATACACGTTGTA 349
 |||||
 Db 126 AAAGAGACATTAAGAGAGAA 146

RESULT 15

CE571358/c

CE571358 590 bp DNA linear GSS 28-SEP-2003
 tigr-gss-dog-17000327517076 dog library Canis familiaris genomic,
 genomic survey sequence.

LOCUS CE571358
 DEFINITION CE571358

ACCESSION CE571358.1 GI:36888139
 VERSION CE571358

KEYWORDS GSS.
 SOURCE Canis familiaris (dog)
 ORGANISM Canis familiaris

REFERENCE Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 1 (bases 1 to 590)
 Kirkenes, E.F., Balina, V., Halpern, A.L., Levy, S., Remington, K.,
 Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
 Venter, J.C.

TITLE The dog genome: survey sequencing and comparative analysis
 JOURNAL Science 301 (5641), 1898-1903 (2003)
 MEDLINE 22873432
 PUBMED 14512627

COMMENT Contact: Kirkenes EF
 The Institute for Genomic Research
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
 Rockville, MD 20850, USA
 Tel: 301-838-0200
 Fax: 301-838-0208
 Email: ekirkenes@tigr.org
 Class: shotgun.

FEATURES
 source Location/Qualifiers

1..590
 /organism="Canis familiaris"
 /mol_type="genomic DNA"
 /strain="Standard Poodle"
 /db_xref="taxon:9615"
 /clone_lib="Dog library"
 /note="Site 1: BstXI; Libraries were prepared from
 peripheral blood"

ORIGIN

Query Match 2.7%; Score 37; DB 9; Length 590;

Best Local Similarity 58.7%; Pred. No. 18;
 Matches 64; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1218 TAGATACATAGAGTATAGAAACATAAAGAGCTTGAAGACCATTCATCAATCCTTAAGGG 1277
 |||||
 Db 326 TACCTAAATGAAACATAAATATATATGAGGTTCAAAAATTTATTACTTCCCAAGGG 267
 |||||
 QY 1278 TCTCTCTTTCTTCGATCATCATCAAGATCATCATCAACCAAGGAA 1326
 |||||
 Db 266 GCTTTTGCATCTGATCTCATCTGTATTATTTCACAAAGATCAAAAA 218
 |||||

Search completed: February 9, 2005, 13:39:56
 Job time : 4291 secs